(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
29 August 2002 (29.08.2002)

PCT

(10) International Publication Number
WO 02/066514 A2

(51) International Patent Classification[7]: C07K 16/00

(21) International Application Number: PCT/EP02/01690

(22) International Filing Date: 18 February 2002 (18.02.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
01103955.9 19 February 2001 (19.02.2001) EP
01108291.4 5 April 2001 (05.04.2001) EP

(71) Applicant *(for all designated States except US)*: **MERCK PATENT GMBH** [DE/DE]; Frankfurter Strasse 250, 6100 Darmstadt (DE).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **GILLIES, Stephen** [US/US]; 159 Sunset Road, Carlisle, MA 01741 (US). **CARR, Francis, J.** [GB/GB]; Birchlea, The Holdings, Balmedie, Aberdeenshire AB23 8XU (GB). **JONES, Tim** [GB/GB]; 27 Brick Row, Babraham, Cambridge CB2 4AJ (GB). **CARTER, Graham** [GB/GB]; Longhill Cottage, By Newmachar, Aberdeenshire AB21 7XB (GB). **HAMILTON, Anita** [GB/GB]; 130 Kings Gate, Aberdeen AB15 4EQ (GB). **WILLIAMS, Stephen** [GB/GB]; Drumgowan Cottage, Auchleven, Insch, Aberdeeshire AB52 6QD (GB). **HANLON, Marian** [GB/GB]; 61 Vinery Road, Cambridge CB1 3DN (GB). **WATKINS, John** [GB/GB]; 62 Cambridge Road, Girton, Cambridge CB3 0PJ (GB). **BAKER, Matthew** [GB/GB]; 8 Saffron Close, Littleport, Ely, Cambridge GB6 1AR (GB). **WAY, Jeffrey, C.** [US/US]; 12 Avon Place, Cambridge, MA 02140 (US).

(74) Common Representative: **MERCK PATENT GMBH**; Frankfurter Strasse 250, 64293 Darmstadt (DE).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/066514 A2

(54) Title: ARTIFICIAL PROTEINS WITH REDUCED IMMUNOGENICITY

(57) Abstract: The invention relates to artificial modified proteins, preferably fusion proteins, having a reduced immunogenicity compared to the parent non-modified molecule when exposed to a species *in vivo*. The invention relates, above all, to novel immunoglobulin fusion proteins which essentially consist of an immunoglobulin molecule or a fragment thereof covalently fused via its C-terminus to the N-terminus of a biologically active non-immunoglobulin molecule, preferably a polypeptide or protein or a biologically active fragment thereof. In a specific embodiment, the invention relates to fusion proteins consisting of an Fc portion of an antibody which is fused as mentioned to the non-immunological target molecule which elicits biological or pharmacological efficacy. The molecules of the invention have amino acid sequences which are altered in one or more amino acid residue positions but have in principal the same biological activity as compared with the non-altered molecules. The changes are made in regions of the molecules which are identified as T-cell epitopes, which contribute to an immune reaction in a living host. Thus, the invention also relates to a novel method of making such fusion proteins by identifying said epitopes comprising calculation of T-cell epitope values for MHC Class II molecule binding sites in a peptide by computer-aided methods.

ARTIFICIAL PROTEINS WITH REDUCED IMMUNOGENICITY

TECHNICAL FIELD OF THE INVENTION

The invention relates to modified artificial proteins, preferably fusion proteins, having a reduced immunogenicity compared to the parent non-modified molecule when exposed to a species *in vivo*. Especially, the invention relates to proteins that are as single component normally not strongly immunogenic, but which have enhanced immunogenicity when attached to a second protein moiety to form an, as a rule, artificial fusion protein. The invention relates, above all, to modified and, thus, novel immunoglobulin (Ig) fusion proteins which essentially consist of an immunoglobulin molecule or a fragment thereof covalently fused via its C-terminus to the N-terminus of a biologically active non-immunoglobulin molecule, preferably a polypeptide or protein or a biologically active fragment thereof. In a specific embodiment, the invention relates to fusion proteins consisting of an Fc portion of an antibody which is fused as mentioned to the non-immunological target molecule which elicits biological or pharmacological efficacy.

The molecules of the invention have amino acid sequences which are altered in one or more amino acid residue positions but have in principal the same biological activity as compared with the non-altered molecules. The changes are made in regions of the molecules which are identified as T-cell epitopes, which contribute to an immune reaction in a living host. Thus, the invention also relates to a novel method for preparing said fusion proteins by identifying said epitopes comprising calculation of T-cell epitope values for MHC Class II molecule binding sites in a peptide by computer-aided methods.

BACKGROUND OF THE INVENTION

Therapeutic fusion proteins are, as a rule, artificial molecules, which are produced to combine known favorable properties of the single components or to create new properties. For example, a fusion protein may contain an immunogenic moiety that causes a normally non-immunogenic fusion partner to become immunogenic. In other cases, each of the components are immunogenic and the fusion molecule has kept this usually undesired property. Finally, it is possible that fusing non or less immunogenic components the fusion product is immunogenic by creating the bonds, especially the junction region.

Fusion proteins of specific interests in this context are immunoconjugates. Immunoconjugates are known since a couple of years and many of them have shown pharmacological efficacy in

vitro and in vivo. Immunoconjugates are chimeric molecules consisting, as a rule, of a portion deriving from an immunoglobulin or a fragment thereof and a target polypeptide or protein which is linked to the immunoglobulin molecule. Originally, immunoconjugates were prepared consisting of a complete antibody and a cytotoxic agent like a cytokine which was fused via its N- terminus to the C-terminus of the constant domain of the immunoglobulin, or alternatively, via its C-terminus to the N-terminus of the variable region of the antibody (see, for example EP 0439 095, WO 92/08495, US 5,349,053, EP 0659 439, EP 0706 799). These chimeric molecules are bi-functional by targeting a specific antigen, for example, on a tumor cell surface by means of the binding sites within the CDRs of the variable domain of the antibody portion or a fragment thereof, and by the simultaneous cytotoxic effect of the cytokine which is coupled to the immunoglobulin and thus can, theoretically, only or predominantly attack the targeted cell. In this context, also tri- and multi-functional immunoconjugates were developed including constructs consisting of sFv-, Fab-, Fab' or F(ab')2 fragments of different antibodies, wherein in each case the targeting function of the immunoglobulin portion was advantageously used.

Another immunoglobulin modification is the use of the Fc region of antibodies. Antibodies comprise two functionally independent parts, a variable domain known as "Fab" , "Fab", "F(ab')2", dependent on the kind of digestion of the molecule, which bind antigen, and a constant domain, known as "Fc" which provides the link to effector functions such as complement or phagocytic cells. The Fc portion of an immunoglobulin has a long plasma half-life, whereas the Fab fragments are short-lived (Capon, *et al.*, Nature 337: 525-531 (1989)).

Therapeutic protein products have been constructed using the Fc domain to provide longer half-life or to incorporate functions such as Fc receptor binding, protein A binding, complement fixation and placental transfer which all reside in the Fc proteins of immunoglobulins. For example, the Fc region of an IgG1 antibody has been fused to the N-terminal end of CD30-L, a molecule which binds CD30 receptors expressed on Hodgkin's Disease tumor cells, anaplastic lymphoma cells, T-cell leukemia cells and other malignant cell types (US 5,480,981). IL-10, an anti-inflammatory and antirejection agent has been fused to murine Fcγ2a in order to increase the cytokine's short circulating half-life (Zheng *et al.*, The Journal of Immunology, 154: 5590-5600 (1995)). Studies have also evaluated the use of tumor necrosis factor receptor linked with the Fc protein of human IgG1 to treat patients with septic shock (Fisher *et al., N.* Engl. J. Med., 334: 1697-1702 (1996); Van Zee *et al.*,

The Journal of Immunology, 156: 2221-2230 (1996)). Fc has also been fused with CD4 receptor to produce a therapeutic protein for treatment of AIDS (see: Capon *et* al., Nature, 337:525-531 (1989)). Principally, Fc can be fused to the target protein or peptide via its C- or N-terminus using the N- and C-terminus of the protein, respectively. A chimer of Fc and TNF and EPO was disclosed in EP 0464 533 (Hoechst/ General Hospital), wherein the N-terminus of Fc was coupled to the C-terminus of the protein (X-Fc). The identical conjunction was selected for leptin-Fc chimers disclosed in WO 97/00319 (SKB) and WO 97/24440 (Genentech). There are a lot of publications and patent applications describing the opposite linkage of Fc-protein chimers (Fc-X), such as Fc-(IL-2), Fc-EPO, Fc-PSMA, Fc-(IL-12), Fc-TNFa, Fc-(GM-CSF), Fc-TNFR, Fc-endostatin, Fc, angiostatin, Fc-gp120, Fc-leptin, Fc-IFNa, Fc-(G-CSF). Examples are WO 96/08570, (Fuji / Merck KGaA), WO 98/28427 (Amgen), WO 99/02709 (Beth Israel Medical Care Center) and WO 99/58662 (Fuji / Merck KGaA). WO 00/24782 (Amgen) discloses a huge number of possible Fc-X conjugates, wherein the linkage between the two partners may be Fc-X or X-Fc. An extensive development of Fc-X molecules was realized by Lexigen / Merck KgaA as disclosed in US 5,541,087, WO 99/43713, WO 99/29732, WO 99/52562, WO 99/53958, WO 00/11033, WO 01/07081, PCT/EP00/10843.Thus, X-Fc and Fc-X molecules which have "lost" their antigen binding sites, as well as molecules, wherein the binding sites und thus their antigen-specific targeting functions are conserved, are of great interest as promising therapeutic proteins and there exists a further need to develop analogue compositions for different clinical application.

Non-natural therapeutic proteins are often particularly immunogenic. For example, Enbrel is a fusion protein consisting of an extracellular domain of a Tumor Necrosis Factor Receptor (TNF-R) fused to an Fc region of an antibody. About 16% of patients treated with Enbrel have been reported to develop antibodies to this fusion protein (*Physician's Desk Reference* [2001] p. 3372). Similarly, a fusion protein consisting of erythropoietin (Epo) and granulocyte / macrophage-colony stimulating factor (GMCSF) was found to be highly immunogenic (Coscarella A, et al. [1998] Cytokine 10:964-9; Coscarella A, Mol Biotechnol. [1998] 10:115-22). When injected into a primate, Epo-GMCSF fusion proteins were found to induce a strong antibody response to the Epo moiety of the fusion protein, resulting in anemia.

Ceredase™ and Cerezyme™ are forms of the lysosomal enzyme glucocerebrosidase used to treat Gaucher's disease; as a result of genetic engineering, glucocerebrosidase is attached to an unusual high-mannose glycosylation. Patients with Gaucher's disease lack glucocerebrosidase in their lysosomes, and as a result the patients' macrophages tend to accumulate lipids and

become foam cells (*The Metabolic and Molecular Bases of Inherited Disease*, 8th Edition [2001] Scriver et al. eds. Chapter 146, "Gaucher Disease." p. 3635-3668). After administration of Ceredase™ or Cerezyme™, the therapeutic protein is bound by mannose receptors on macrophages, endocytosed, and trafficked through the endosomes to the lysosome, which is its proper location. Patients treated with Ceredase often develop antibodies to glucocerebrosidase (Pastores GM, et al., Blood [1993] 82:408-16; *Physicians' Desk Reference* [2001] p. 1325-1326). Such antibodies can interfere with treatment (Brady RO, et al., Pediatrics. [1997] 100(6):E11.). In a Phase I clinical trial using an antibody-cytokine fusion protein, several patients developed antibody responses to the therapeutic fusion protein. In this case, the antibody moiety was a humanized form of the 14.18 antibody, and the cytokine was interleukin-2 (IL-2). Many of the reactive patients' sera included significant levels of anti-idiotype antibodies.

Therapeutic use of a number of peptides, polypeptides and proteins is curtailed because of their immunogenicity in mammals, especially humans. For example, when murine antibodies are administered to patients who are not immunosuppressed, a majority of such patients exhibit an immune reaction to the introduced foreign material by making human anti-murine antibodies (HAMA) (e.g. Schroff, R. W. et al (1985) *Cancer Res.* 45: 879-885; Shawler, D.L. et al (1985) *J. Immunol.* 135: 1530-1535). There are two serious consequences. First, the patient's anti-murine antibody may bind and clear the therapeutic antibody or immunoconjugate before it has a chance to bind, for example to a tumor, and perform its therapeutic function. Second, the patient may develop an allergic sensitivity to the murine antibody and be at risk of anaphylactic shock upon any future exposure to murine immunoglobulin.

Several techniques have been employed to address the HAMA problem and thus enable the use in humans of therapeutic monoclonal antibodies (see, for example, WO89/09622, EP0239400, EP0438310, WO91/09967). These recombinant DNA approaches have generally reduced the mouse genetic information in the final antibody construct whilst increasing the human genetic information in the final construct. Notwithstanding, the resultant "humanized" antibodies have, in several cases, still elicited an immune response in patients (Issacs J.D. (1990) *Sem. Immunol.* 2: 449, 456; Rebello, P.R. et al (1999) *Transplantation* 68: 1417-1420).

A common aspect of these methodologies has been the introduction into the therapeutic antibody, usually of rodent origin, of amino acid residues, even significant tracts of amino acid residue sequences, identical to those present in human antibody proteins. For antibodies, this process is possible owing to the relatively high degree of structural (and functional) conservatism among antibody molecules of different species. For potentially therapeutic peptides, polypeptides and proteins, however, where no structural homologue may exist in the host species (e.g., human) for the therapeutic protein, such processes are not applicable. Furthermore, these methods have assumed that the general introduction of a human amino acid residue sequence will render the re-modeled antibody non-immunogenic. It is known, however, that certain short peptide sequences ("T-cell epitopes") can be released during the degradation of peptides, polypeptides or proteins within cells and subsequently be presented by molecules of the major histocompatability complex (MHC) in order to trigger the activation of T-cells. For peptides presented by MHC Class II, such activation of T-cells can then give rise to an antibody response by direct stimulation of B-cells to produce such antibodies. Accordingly, it would be desirable to eliminate potential T-cell epitopes from a peptide, polypeptide or a protein. Even proteins of human origin and with the same amino acid sequences as occur within humans can still induce an immune response in humans. Notable examples include therapeutic use of granulocyte-macrophage colony stimulating factor (Wadhwa, M. et al (1999) Clin. Cancer Res. 5: 1353-1361) and interferon alpha 2 (Russo, D. et al (1996) *Bri. J. Haem.* 94: 300-305; Stein, R. et al (1988) *New Engl. J. Med.* 318: 1409-1413).

During the last couple of years several techniques were published which suggest solutions for rendering antibodies and target proteins having different biological functions non- or at least less immunogenic. Examples are: WO 92/10755 and WO 96/40792 (Novo Nordisk), EP 0519 596 (Merck & Co.), EP 0699 755(Centro de Immunologia Moelcular), WO 98/52976 and WO 98/59244 and WO 00/34317 (Biovation Ltd.).

The general methods disclosed in the prior art and regarding the elimination of T-cell epitopes from proteins (e.g. WO 98/52976, WO 00/34317) comprise the following steps:

(a) Determining the amino acid sequence of the polypeptide or part thereof

(b) Identifying one or more potential T-cell epitopes within the amino acid sequence of the protein by any method including determination of the binding of the peptides to MHC molecules using *in vitro* or *in silico* techniques or biological assays.

(c) Designing new sequence variants with one or more amino acids within the identified potential T-cell epitopes modified in such a way to substantially reduce or eliminate the activity of the T-cell epitope as determined by the binding of the peptides to MHC molecules using *in vitro* or *in silico* techniques or biological assays. Such sequence variants are created in such a way to avoid creation of new potential T-cell epitopes by the sequence variations unless such new potential T-cell epitopes are, in turn, modified in such a way to substantially reduce or eliminate the activity of the T-cell epitope.

(d) Constructing such sequence variants by recombinant DNA techniques and testing said variants in order to identify one or more variants with desirable properties.

Other techniques exploiting soluble complexes of recombinant MHC molecules in combination with synthetic peptides and able to bind to T-cell clones from peripheral blood samples from human or experimental animal subjects have been used in the art [Kern, F. et al (1998) *Nature Medicine* 4:975-978; Kwok, W.W. et al (2001) *TRENDS in Immunology* 22: 583-588] and may also be exploited in an epitope identification strategy.

The potential T-cell epitopes are generally defined as any amino acid residue sequence with the ability to bind to MHC Class II molecules. Such potential T-cell epitopes can be measured to establish MHC binding. In the general understanding the term "T-cell epitope" is an epitope which when bound to MHC molecules can be recognized by the T-cell receptor, and which can, at least in principle, cause the activation of these T-cells. It is, however, usually understood that certain peptides which are found to bind to MHC Class II molecules may be retained in a protein sequence because such peptides are tolerated by the immune within the organism into which the final protein is administered.

The invention is conceived to overcome the practical reality that soluble proteins introduced into an autologous host with therapeutic intent, can trigger an immune response resulting in development of host antibodies that bind to the soluble protein. One example amongst others is interferon alpha 2 to which a proportion of human patients make antibodies despite the fact that this protein is produced endogenously [Russo, D. et al (1996) *Brit. J. Haem.* 94: 300-305; Stein, R. et al (1988) *New Engl. J. Med.* 318: 1409-1413]

MHC Class II molecules are a group of highly polymorphic proteins which play a central role in helper T-cell selection and activation. The human leukocyte antigen group DR (HLA-DR) are the predominant isotype of this group of proteins and the major focus of the present invention. However, isotypes HLA-DQ and HLA-DP perform similar functions, hence the present invention is equally applicable to these. MHC HLA-DR molecules are homo-dimers where each "half" is a hetero-dimer consisting of α and ß chains. Each hetero-dimer possesses a ligand binding domain which binds to peptides varying between 9 and 20 amino acids in length, although the binding groove can accommodate a maximum of 9 - 11 amino acids. The ligand binding domain is comprised of amino acids 1 to 85 of the α chain, and amino acids 1 to 94 of the ß chain. DQ molecules have recently been shown to have an homologous structure and the DP family proteins are also expected to be very similar. In humans approximately 70 different allotypes of the DR isotype are known, for DQ there are 30 different allotypes and for DP 47 different allotypes are known. Each individual bears two to four DR alleles, two DQ and two DP alleles. The structure of a number of DR molecules has been solved and such structures point to an open-ended peptide binding groove with a number of hydrophobic pockets which engage hydrophobic residues (pocket residues) of the peptide [Brown et al *Nature* (1993) 364: 33; Stern et al (1994) *Nature* 368: 215]. Polymorphism identifying the different allotypes of class II molecule contributes to a wide diversity of different binding surfaces for peptides within the peptide binding grove and at the population level ensures maximal flexibility with regard to the ability to recognize foreign proteins and mount an immune response to pathogenic organisms.

There is a considerable amount of polymorphism within the ligand binding domain with distinct "families" within different geographical populations and ethnic groups. This polymorphism affects the binding characteristics of the peptide binding domain, thus different "families" of DR molecules will have specificities for peptides with different sequence properties, although there may be some overlap. This specificity determines recognition of Th-cell epitopes (Class II T-cell response) which are ultimately responsible for driving the antibody response to B-cell epitopes present on the same protein from which the Th-cell epitope is derived. Thus, the immune response to a protein in an individual is heavily influenced by T-cell epitope recognition which is a function of the peptide binding specificity of that individual's HLA-DR allotype. Therefore, in order to identify T-cell epitopes within a protein or peptide in the context of a global population, it is desirable to consider the binding

properties of as diverse a set of HLA-DR allotypes as possible, thus covering as high a percentage of the world population as possible.

A principal factor in the induction of an immune response is the presence within the protein of peptides that can stimulate the activity of T-cell via presentation on MHC class II molecules. In order to eliminate or reduce immunogenicity, it is thus desirable to identify and remove T-cell epitopes from the protein.

According to the above-cited methods and related processes several biological molecules, basically usual target proteins and antibodies have been prepared which reveal reduced immunogenicity and allergenicity. Examples are: WO 99/55369 (SKB), WO 99/40198 and WO 96/21016 (Leuven Research & Development VZW), WO 00/08196 (Duke University), WO 96/21036 (Chiron Viragen), WO 97/31025 (Chiron Corp.), WO 98/30706 (Alliance Pharmaceutical Corp.).

In all these applications cited above single proteins or antibodies eliciting a lower immune response were disclosed; there is no hint that fusion proteins, above all immunoglobulin fusion proteins were completely or partially de-immunized, especially by reducing the number of T-cell epitopes within the sequence of said molecules by means of partially computational methods. In WO 97/24137 (Tannox Biosystems Inc.) a IFNα-Fc chimer is disclosed which contains a non-immunogenic linker molecule between the N-terminus of the Fc portion and the C-terminus of IFNα.

Therefore, it is still a need to provide for biological molecules, such as immunoconjugates, which are not or less immunogenic. Above all, it is of specific interest to provide for Fc-conjugates, preferably Fc-X chimers, wherein X is a selected protein or polypeptide of therapeutic interest.

SUMMARY OF THE INVENTION

The present invention relates to four general aspects:

(a) a novel application of the details of the immune response mechanism to situations involving fusion proteins and other artificial proteins, to help determine when an engineered or novel protein is likely to be immunogenic and therefore when application of a deimmunization methodology is appropriate,

(b) novel biologically active artificial proteins to be administered especially to humans and in particular for therapeutic use,

(c) a method of designing improved, less immunogenic artificial proteins that normally have enhanced immunogenicity, the method comprising identification of one or more candidate T-cell epitopes in the artificial protein and introducing a mutation that removes one or more T-cell epitopes, and

(d) a convenient and effective computational method for the identification and calculation of T-cell epitopes for a globally diverse number of MHC Class II molecules and, based on this knowledge, for designing and constructing new sequence variants of biological molecules with improved properties. Once T-cell epitopes have been identified in an artificial protein, they are removed by mutation as described in (c).

Artificial proteins that have a component capable of binding to a surface receptor on a cell of the immune system are, in general, particularly immunogenic. Artificial proteins that are immunogenic as a consequence of having a moiety that binds to an immune cell surface receptor are particularly good candidates for the methods of the invention for reducing immunogenicity.

Without wishing to be bound by theory, Figures 1-6 present diagrams of artificial proteins containing moieties that bind to immune cell surface receptors such as an Fc receptor, a cytokine receptor, or an oligosaccharide receptor.

One method of the invention consists of the steps of identifying artificial proteins that contain moieties that bind to an immune cell surface receptor, which may be done by sequence inspection, identifying candidate T-cell epitopes in the artificial protein, designing mutant derivatives of the artificial protein in which the number of T-cell epitopes is reduced, producing one or more mutant derivatives, testing the mutant derivatives for activity and optionally other desired properties, and choosing a mutant derivative that has an optimal balance of reduced T-cell epitopes, retained activity, and optionally other retained desired properties. Other desired properties may include, but are not limited to, pharmacokinetic properties and protein expression and assembly characteristics.

Artificial proteins that tend to form aggregates are a second category of proteins that can be improved by the methods of the invention.

One class of artificial proteins that can particularly be improved by the invention are Ig fusion proteins, such as fusion proteins comprising an entire antibody, as well as Fc-X and X-Fc fusion proteins. In particular, immunoglobulin fusion proteins comprising a functional Fc receptor binding site can be particularly improved by methods of the invention.

The invention provides improved forms of such antibody fusion proteins, which include fusion proteins comprising V regions that recognize tumor-specific antigens, other tissue-specific antigens, or other disease-specific antigens. In one preferred embodiment, each of these antibodies is fused to a cytokine, such as IL-2.

For example, the invention provides fusion proteins comprising the tumor-directed anti-EpCAM antibody KS 1/4 and anti-GD2 antibody 14.18, in which the V regions of the antibody contain mutations that remove T-cell epitopes.

In a distinct embodiment, the moiety that is fused to the antibody moiety is mutated such that T-cell epitopes are removed. For example, the invention discloses an antibody-IL-2 fusion protein in which the IL-2 moiety has been altered to remove T-cell epitopes.

A second general class of Ig fusion proteins that can be significantly improved are the Fc-X and X-Fc fusion proteins. Without wishing to be bound by theory, it is thought that these proteins are particularly immunogenic because the Fc receptor binding site, which is normally somewhat sterically blocked by the light chain in an intact antibody, is exposed. In any case, it has been empirically established that an Fc fusion protein can be more immunogenic than the fusion partner by itself (WO01/07081).

Another class of immunogenic fusion proteins are proteins that are fused to a cytokine. Without wishing to be bound by theory, it may be that these proteins are particularly immunogenic because when the fusion partner protein binds to an immune cell, for example a cell bearing an antibody that recognizes the fusion partner protein, the cytokine stimulates the cell in some way (see Figure 4).

A class of artificial proteins that are particularly immunogenic are normal proteins that contain an inappropriate oligosaccharide. For example, a protein containing an oligosaccharide that is bound by a specific receptor on an immune cell is often found to be immunogenic. For example, a protein, preferably a protein such as beta-glucocerebrosidase that can be used to treat a lysosomal storage disorder, contains a high mannose oligosaccharide. Such an immunogenic protein shows significantly reduced immunogenicity when modified according to the invention.

The invention provides less immunogenic forms of the following protein moieties that are incorporated into otherwise immunogenic fusion proteins such as: erythropoietin, leptin, keratinocyte growth factor, G-CSF GM-CSF, IL-1R antagonist, sTNFR, TNF inhibitor, sTNFR-Fc (Enbrel®), BNTF, CNTF, members of the interferon family, hGH, ß-glucocerebrosidase. All these biologically active protein moieties listed above derive from well known non-modified (parent) protein moieties according to the invention.

The modified proteins according to the invention may be produced by the method indicated in Section "Detailed Description of the Invention". The method includes a novel method for identification T-cell epitopes by computational means. This method step is preferred according to the invention and described in more detail in EXAMPLE 1.

The invention discloses and claims as preferred embodiments of the invention altered or modified fusion proteins derived from parent fusion proteins, said parent fusion protein essentially consisting of an immunoglobulin molecule or a fragment thereof and a non-immunoglobulin target polypeptide (X), which is linked preferably by its N-terminal to the C-terminal of the immunoglobulin molecule or a fragment thereof, wherein the altered fusion protein has an amino acid sequence different from that of said parent fusion protein and exhibits reduced immunogenicity relative to the parent fusion protein when exposed to the immune system of a given species, that is preferably human.

The strategies that are used in practice according to the invention to reduce the immunogenicity of an immunogenic fusion protein are illustrated in detail for the antibody-cytokine fusions. These general strategies include:

- Examining the amino acid sequences in the fusion protein and prioritizing them with respect to likely immunogenicity, based on the expected presence and abundance of the sequences during negative selection of T-cells in the thymus. For example, completely non-self epitopes are identified, and are the highest priority for removal of T-cell epitopes by mutation. The lowest priority for removal of T-cell epitopes are sequences that are present in abundant serum proteins, such as antibody constant regions or sequences that are found in un-rearranged antibody V regions. An intermediate priority for removal of T-cell epitopes by mutation are self sequences that are found in low abundance proteins, such as cytokines. Without wishing to be bound by theory, it is expected that low

abundance proteins may not be present in the thymus in high enough amounts to promote negative T-cell selection, and may thus be recognized as non-self T-cell epitopes.

- When a region is chosen for removal of T-cell epitopes by mutation, it is compared with naturally occurring human sequences found in abundant proteins. Mutations are introduced to make any non-self sequences more similar to self sequences. For example, to reduce the immunogenicity of a mouse V region, the sequence is compared to un-rearranged human V regions and the most closely related sequence is found. "Veneering" changes are introduced, in which some amino acids are converted from mouse to human. This has the effect of converting some non-self T-cell epitopes into self T-cell epitopes, a method for reducing immunogenicity disclosed by US 5712120, and also has the effect of removing some B cell epitopes. However, it is still necessary to remove T-cell epitopes that derive from hypervariable region sequences.
- To remove most or all of the remaining T-cell epitopes, mutations are introduced that, by the computer-based criteria defined above, prevent binding of a peptide into a groove of an MHC Class II molecule. In the case of antibody V regions, it is preferable to introduce mutations that lie outside the CDRs themselves, to avoid interfering with antigen binding.
- In the case of fusion proteins of any type, it is generally the case that the fusion junction will contain non-self T-cell epitopes. These T-cell epitopes may be also removed by mutation.

As a specific embodiment the invention includes chimeric immunoglobulins or fragments thereof wherein the reduced immunogenicity, the reduced number of T-cell epitopes or the reduced number of peptides binding to MHC class II molecules is located to the target polypeptide portion X as well as to the immunoglobulin portion or fragments thereof of the altered fusion protein.

The invention includes also chimeric immunoglobulins as defined according to the invention wherein the immunoglobulin molecule and the non-immunoglobulin target polypeptide (X) are fused via a linker molecule (L). As a specific embodiment of the invention this linker molecule itself has no or lower immunogenicity. Thus, the invention may include immunoconjugates, wherein the linker molecule alone is de-immunized . Linker molecules which have reduced or no immunogenicity are known in the art or can be prepared by known methods or by the method according to the invention. The invention also includes such

immunoglobulin fusion proteins, wherein the immunoglobulin portion as well as the target protein (X) portion of the fusion molecule and optionally the linker molecule and the junction region (see below) are immunogenicly modified. Alternatively, only one or more but not all portions of the molecule are modified according to the invention.

The invention relates, furthermore, to above-said immunoconjugates which may derive, in principal, from all immunoglobulin classes; however IgG is preferred. It is an object of the invention to provide such chimeric immunoglobulins which derive from IgG1, IgG2, IgG3 and IgG4. IgG1 and IgG2 immunoglobulins are preferred; IgG2 immunoglobulins are most preferred.

Since it has been shown that even recombinant proteins of human origin and humanized antibodies may elicit an undesirable immune response in humans it is object of this invention to provide fusion proteins wherein the immunoglobulin portion as well as the target polypeptide portion (X) may be selected from non-human as well as from human origin. Since humanized or human-derived molecules have, as a rule, a less number of T-cell epitopes, such molecules are preferred for de-immunization, because a less number of amino acid residues has to be modified.

Immunoconjugates (immunoglobulin (Ig) fusion proteins) according to the invention include also fragments of antibodies like sFv, Fab, Fab', F(ab')2 and Fc. It is a specific and preferred object of the invention to provide said above- and below-defined fusion proteins, wherein the immunoglobulin portion is a Fc domain of an antibody, preferably an IgG1 or IgG2 antibody. Fc-X molecules according to the invention which have reduced affinity to Fc receptors are a preferred object of the invention. Fc molecules having a reduced affinity to Fc receptors are well known in the art and can be prepared by modifying the amino acid sequence of the Fc domain (e.g. WO 99/43713).

In detail, the invention refers to:

- an immunogenicly modified fusion protein derived from a parent fusion protein, essentially consisting of a first protein / polypeptide and a second protein / polypeptide, wherein the first protein is an immunoglobulin molecule or a fragment thereof and the second protein / polypetide is non-immunoglobulin target polypeptide (X) each linked to the other directly or by a linker molecule; said modified fusion protein having an amino

acid sequence different from that of said parent fusion protein and exhibiting reduced immunogenicity by a reduced number of T-cell epitopes within its amino acid sequence relative to the parent fusion protein when exposed to the immune system of a given species;

- a corresponding fusion protein, wherein said T-cell epitopes are peptide sequences able to bind to MCH class II molecule binding groups;
- a corresponding fusion protein, wherein the target polypeptide (X) is linked by its N-terminal to the C-terminal of the immunoglobulin moiety;
- a correspondingly modified fusion protein, wherein the given species is a human;
- a corresponding fusion protein, wherein the fusion components are fused via a linker molecule L;
- a modified fusion protein according to claim 4, wherein said linker molecule L is non-immunogenic or less immunogenic;
- a corresponding fusion protein, wherein the junction region represented by the C-terminal region of the immunoglobulin portion and the N-terminal region of the non-immunoglobulin target polypeptide (X) has no or a reduced number of T-cell epitopes;
- a corresponding fusion protein, wherein the immunoglobulin portion or a fragment thereof or the target polypeptide (X) portion is less immunogenic;
- a corresponding fusion protein, wherein said immunoglobulin molecule or fragment thereof is IgG1 or IgG2;
- a corresponding fusion protein, wherein said immunoglobulin fragment is a Fc portion, wherein, preferably, said Fc portion has a reduced affinity to Fc receptors;
- an immunogenicly modified fusion protein according having the formula

$$Fc - L_n - X$$

wherein

Fc is the Fc portion of an immunoglobulin molecule (antibody),

X is a non-immunoglobulin target polypeptide

L is a linker peptide,

n = 0 or 1, and

wherein X and / or L comprises amino acid residue modifications which elicit a reduced immunogenicity compared to the parent molecule.

Preferred embodiments of these immunogenicly modified Fc fusion molecules are:

$Fc - X^m$, wherein X is modified only,

$Fc - L^m - X^m$, wherein X and L are modified to have a reduced immunogenicity,

$Fc - X^m$, wherein X and the junction region between Fc and X are modified,

$Fc - L^m - X^m$, wherein X and L and the junction regions between Fc , X and L are modified;

- a corresponding Fc – (L) – X fusion protein wherein at least X is immunogenicly modified;
- an immunogenicly modified fusion protein having the formula

$$A - L_n - X$$

wherein

A is a whole antibody or its sFv, Fab, Fab', $F(ab')_2$ fragments

X is a non-immunoglobulin target polypeptide

L is a linker peptide,

n = 0 or 1, and

wherein A and / or X and / or L comprise amino acid residue modifications which elicit a reduced immunogenicity compared to the parent molecule;

Preferred embodiments of these immunogenicly modified fusion molecules are:

$A - X^m$, wherein X is modified only, optionally the A-X junction region,

$A^m - X^m$, wherein A and X are modified, optionally their junction region,

$A - L^m - X^m$, wherein X and L are modified only to have a reduced immunogenicity,

$A^m - X$, wherein A has reduced immunogenicity only, optionally the A-X junction region,

$A^m - L^m - X^m$, wherein A, L and X are immunogenicly modified, optionally the A-L-X junction regions;

- a corresponding A – (L) – X fusion molecule, wherein at least X or A is immunogenicly modified;
- a correponding fusion protein, wherein A is selected from the group:

anti- EGF receptor (HER1) antibodies

anti- HER2 antibodies

anti- CDx antibodies, wherein x is an integer from 1 - 25

anti- cytokine receptor antibodies

anti- 17-1A antibodies,

anti- KSA antibodies

anti-GP IIb/IIIa antibodies

anti-integrin receptor antibodies

anti VEGF receptor antibodies;

- a correpondingly fusion protein, wherein the antibody is selected from the group:

monoclonal antibody 225 and derivatives,

monoclonal antibody 425 and derivatives

monoclonal antibody KS 1/4 and derivatives

monoclonal antibody 14.18 and derivatives

monoclonal antibody 4D5 / HER2 (Herceptin®) and derivatives

monoclonal antibody 17-1A and derivatives

monoclonal anti-CD3 antibodies

monoclonal antibody 7E3 and derivatives

monoclonal antibodies LM609, P1F6 and 14D9.F8 and derivatives

monoclonal antibody DC-101 and derivatives

monoclonal anti-Il-2R antibody (Zenapax®) and derivatives

- a corresponding fusion protein, wherein the target polypeptide X is selected from the group:

cytokines, integrin inhibitors, soluble cytokine receptors, glycoproteins, hormones, glycoprotein hormones, leptin, growth hormones, growth factors, antihemophilic factors, antigens, cytokine receptor antagonists;

- a corresponding fusion protein, wherein the target polypeptide X is selected from the group:

IL-2, G-CSF, GM-CSF, EPO, TPO, members of the interferon family, TNFα, soluble TNF receptor, IL-12, IL-8, factor VIII, FGF, TGF, EGF, VEGF, PMSA, IGF, insulin, RGD-peptides, endostatin, angiostatin, BDNF, CNTF, protein c, factor VIII and IX, and and biologically active fragments thereof;

- a more specidied corresponding fusion protein selected from the group:

MAb KS 1/4 – IL2, MAb 14.18 – IL2

MAb 425 – IL2, MAb c425 – IL2, MAb h425 – IL2, MAb 425 – TNFa

MAb 225 – IL2, MAb c225 – IL2

MAb 4D5 – IL2, MAb DC101 – Il2, MAb LM609 – IL2,

Fc – IL2 , Fc – TNFa, Fc – G-CSF, Fc – EPO, Fc – Leptin, Fc – KGF,

Fc – BFNF, FC – ß-Cerebrosidase, Fc – TPO, Fc – GM-CSF;

- an immunogenicly modified artificial protein selected from the group:
  (i) Y – (L) – X, wherein Y is a cytokine and X, (L) is a molecule as defined above,
  (ii) P – (L) – X, wherein P is a protein with unusual glycosylation moieties and X, (L) is a molecule as defined above,
  (iii) A – (L) – X, wherein A, X (L) is a molecule as defined above,
  derived from a parent artificial protein having an amino acid sequence which is different from that of said parent artificial protein and exhibits reduced immunogenicity by a reduced number of T-cell epitopes relative to the parent fusion protein when exposed to the immune system of a given species, wherein said T-cell epitopes are peptide sequences able to bind to MCH class II molecule binding groups obtainable or obtained by a method as specified in this invention;
- a DNA sequence encoding any fusion protein as specified above and below;
- a DNA sequence encoding a corresponding fusion protein, comprising
  (i) a signal sequence
  (ii) a DNA sequence encoding all domains or a Fc, sFV, Fab, Fab' or F(ab')2 domain of an IgG1, IgG2 or IgG3 antibody, and
  (ii) a DNA sequence encoding the polypeptide (X), and optionally
  (iii) a DNA sequence encoding the linker molecule;
- an expression vector comprising a corresponding DNA sequence;
- a pharmaceutical composition comprising a fusion protein as specified above and below, optionally together with a suitable carrier, excipient or diluent or another therapeutically effective drug, such as chemotherapeutics or cytotoxic drugs;
- a method for preparing an immunogenicly modified fusion protein as specified comprising the steps:
  (i) determining the amino acid sequence of the parent fusion protein or part thereof;
  (ii) identifying one or more potential T-cell epitopes within the amino acid sequence of the fusion protein by any method including determination of the binding of the peptides to MHC molecules using *in vitro* or *in silico* techniques or biological assays, (iii) designing new sequence variants by alteration of at least one amino acid residue within the originally identified T-cell epitope sequences, said variants are modified in such a way to substantially reduce or eliminate the activity or number of the T-cell epitope sequences and / or the number of MHC allotypes able to bind peptides derived from said biological molecule as determined by the binding of the peptides to MHC molecules using *in vitro* or

in silico techniques or biological assays or by binding of peptide-MHC complexes to T-cells, (iv) constructing such sequence variants by recombinant DNA techniques and testing said variants in order to identify one or more variants with desirable properties, and (v) optionally repeating steps (ii) – (iv), characterized in that the identification of T-cell epitope sequences according to step (ii) is achieved by

(a) selecting a region of the peptide having a known amino acid residue sequence;

(b) sequentially sampling overlapping amino acid residue segments of predetermined uniform size and constituted by at least three amino acid residues from the selected region; (c) calculating MHC Class II molecule binding score for each said sampled segment by summing assigned values for each hydrophobic amino acid residue side chain present in said sampled amino acid residue segment; and (d) identifying at least one of said segments suitable for modification, based on the calculated MHC Class II molecule binding score for that segment, to change overall MHC Class II binding score for the peptide without substantially the reducing therapeutic utility of the peptide;

- a corresponding method, wherein step (c) is carried out by using a Böhm scoring function modified to include 12-6 van der Waal's ligand-protein energy repulsive term and ligand conformational energy term by (1) providing a first data base of MHC Class II molecule models; (2) providing a second data base of allowed peptide backbones for said MHC Class II molecule models; (3) selecting a model from said first data base; (4) selecting an allowed peptide backbone from said second data base; (5) identifying amino acid residue side chains present in each sampled segment; (6) determining the binding affinity value for all side chains present in each sampled segment; and optionally (7) repeating steps (1) through (5) for each said model and each said backbone;
- a corresponding method, wherein the sampled amino acid residue segment is constituted by 13 amino acid residues and / or consecutive sampled amino acid residue segments overlap by one to five amino acid residues;
- a corresponding method, wherein 1 – 9 amino acid residues, preferably one amino acid residue, in any of the originally present T-cell epitope sequences (is) are altered;
- a corresponding method, wherein the alteration of the amino acid residues is substitution, deletion or addition of originally present amino acid(s) residue(s) by other amino acid residue(s) at specific position(s);
- a corresponding method, wherein additionally further alteration by substitution, deletion or addition is conducted to restore biological activity of said biological molecule.

The polypeptides according to the invention include also antigens, like PMSA and others. Antigens which elicit a not desired and too strong immune response can be modified according to the method of the invention and result in antigens which have a reduced immunogenicity which is however strong enough for using the antigen e.g. as vaccine. The invention includes also variants and other modification of a specific polypeptide, protein, fusion protein, immunoglobulin or immunoconjugate which have in principal the same biological activity and a similar (reduced) immunogenicity. All proteins mentioned above are well known and described in the art or are already commercially available. Most of them are known to have a proved therapeutic benefit. The leader or signal sequences and linker sequences may be optional.

Preparing the fusion protein by linking the immunoglobulin component by its C-terminus or its fragment to the N-terminus of the non-immunoglobulin target polypeptide (X), optionally via the linker molecule according to step (ii) as described above, is carried out by:

(i) preparing a gene construct comprising a DNA sequence encoding the polypeptide X, a DNA sequence encoding the immunoglobulin molecule or fragments thereof [sFv, Fab, Fab', $F(ab')_2$, Fc], and optionally the DNA sequence of a the linker molecule, and

(ii) expressing the gene construct by an expression system.

The immunoconjugates according to the present invention reveal enhanced properties. Thus decreased protein degradation, increased stability and enhanced serum circulation half-life can be measured as well as a distinctly reduced immunogenicity and / or allergenicity. Surprisingly, the reduced immunogenicity leads in many cases to a further increase of half-life, especially in cases where Fc-X molecules according to the invention are used. The reduced immunogenicity makes the fusion proteins according to the invention more tolerable for a given species compared to the non-modified fusion proteins and, therefore, can be administered in higher dosages if necessary.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 illustrates one of the mechanisms by which fusion proteins displays enhanced immunogenicity. Figure 1a shows a protein ("X") fused to an Fc moiety binding to a cell bearing an Fc receptor. Figure 1b shows the fusion protein being processed such that the "X"

moiety is preferentially degraded. Figure 1c shows a peptide remnant of "X" being presented by an MHC molecule to a T-cell.

Figure 2 shows a mechanism of enhanced immunogenicity for a fusion protein comprising an Fc moiety and a second moiety. Figure 2a shows the binding of the fusion protein to a B cell that expresses an antibody specific for "X" on its surface. The fusion protein is bound by both the specific antibody and by Fc receptors that are not already bound by the antibody.

Figure 3 illustrates a second mechanism by which fusion proteins displays enhanced immunogenicity. Figure 1a shows a protein ("X") fused to an cytokine moiety binding to a B cell with a surface-bound antibody. Figure 3b shows the fusion protein being processed. Figure 3c shows a peptide remnant of "X" being presented by an MHC molecule to a T-cell, at the same time that additional X-cytokine fusion protein is bound to the surface of the B cell.

Figure 4 illustrates another mechanism by which an engineered protein displays enhanced immunogenicity. In this case, a cytokine-X fusion protein directly activates a B cell. The B cell synthesizes a specific antibody to X, which increases the local concentration of the cytokine in the neighborhood of the B cell.

Figure 5 illustrates another mechanism by which an engineered protein displays enhanced immunogenicity. Figure 5a shows the binding of a protein bearing a glycosylation moiety to a specific cell-surface receptor for that glycosylation moiety on an immune cell. Figure 5b shows the uptake and degradation of the glycosylated protein. Figure 5c shows the presentation of a peptide remnant of the glycosylated protein to a T-cell via an MHC molecule.

Figure 6 shows a mechanism by which an antibody-cytokine fusion protein displays enhanced immunogenicity. Figure 6a shows the binding of the antibody-cytokine fusion protein to a B cell that expresses an antibody specific for the CDRs of the antibody-cytokine fusion protein. The fusion protein is bound by both the specific antibody and by Fc receptors that are not already bound by the antibody. Figure 6b shows the fusion protein being processed. Figure 6c shows a peptide remnant of the CDRs being presented by an MHC molecule to a T-cell, at the same time that additional antibody-cytokine fusion protein is bound to the surface of the B cell.

DETAILED DESCRIPTION OF THE INVENTION

The term "T-cell epitope" means according to the understanding of this invention an amino acid sequence which is able to bind with reasonable efficiency MHC class II molecules (or their equivalent in a non-human species), able to stimulate T-cells and / or also to bind (without necessarily measurably activating) T-cells in complex with MHC class II.

The term "peptide" as used herein and in the appended claims, is a compound that includes two or more amino acids. The amino acids are linked together by a peptide bond (defined herein below). There are 20 different naturally occurring amino acids involved in the biological production of peptides, and any number of them may be linked in any order to form a peptide chain or ring. The naturally occurring amino acids employed in the biological production of peptides all have the L-configuration. Synthetic peptides can be prepared employing conventional synthetic methods, utilizing L-amino acids, D-amino acids, or various combinations of amino acids of the two different configurations. Some peptides contain only a few amino acid units. Short peptides, e.g., having less than ten amino acid units, are sometimes referred to as "oligopeptides". Other peptides contain a large number of amino acid residues, e.g. up to 100 or more, and are referred to as "polypeptides". By convention, a "polypeptide" may be considered as any peptide chain containing three or more amino acids, whereas a "oligopeptide" is usually considered as a particular type of "short" polypeptide. Thus, as used herein, it is understood that any reference to a "polypeptide" also includes an oligopeptide. Further, any reference to a "peptide" includes polypeptides, oligopeptides, and proteins. Each different arrangement of amino acids forms different polypeptides or proteins. The number of polypeptides–and hence the number of different proteins–that can be formed is practically unlimited.

The term "less or reduced immunogenic(ity)" used before and thereafter is a relative term and relates to the immunogenicity of the respective original source molecule when exposed in vivo to the same type of species compared with the molecule modified according to the invention.

The term "modified protein" as used according to this invention describes a protein which has reduced number of T-cell epitopes and elicits therefore a reduced immunogenicity relative to the parent protein when exposed to the immune system of a given species.

The term "non-modified protein" as used according to this invention describes the "parent" protein as compared to the "modified protein" and has a larger number of T- cell epitopes and, therefore, an enhanced immunogenicity relative to the modified protein when exposed to the immune system of a given species.

The term "biologically active protein" as used here and in the claims includes according to the invention polypetides, proteins, immunoglobulins such as antibodies, antibody fragments, fusion proteins, enzymes, antigens and so on, if not defined otherwise, which elicit a biological and / or therapeutic effect.

The term "cytokine" is used herein to describe proteins, analogs thereof, and fragments thereof which are produced and excreted by a cell, and which elicit a specific response in a cell which has a receptor for that cytokine. Preferably, cytokines include interleukins such as interleukin-2 (IL-2), hematopoietic factors such as granulocyte-macrophage colony stimulating factor (GM-CSF), tumor necrosis factor (TNF) such as TNFa, and lymphokines such as lymphotoxin. Preferably, the antibody-cytokine fusion protein of the present invention displays cytokine biological activity. In principal, the inventions encompasses all cytokines as recently classified according to their receptor code (Inglot, 1997, Archivum Immunologiae et Therapiae Experimentalis, 45: 353).

The phrase "single chain Fv" or "scFv" refers to an antibody in which the heavy chain and the light chain of a traditional two chain antibody have been joined to form one chain. Typically, a linker peptide is inserted between the two chains to allow for proper folding and creation of an active binding site.

The term "Fc region" or "Fc domain" as used in this invention is understood to mean the carboxyl terminal portion of an immunoglobulin heavy chain constant region, or an analog or portion thereof capable of binding an Fc receptor. As is known, each immunoglobulin heavy chain constant region comprises four or five domains. The domains are named sequentially as follows: CH1 -hinge-CH2-CH3(-CH4). CH4 is present in IgM, which has no hinge region. The immunoglobulin heavy chain constant region useful in the practice of the invention preferably comprises an immunoglobulin hinge region, and preferably also includes a CH3 domain. The immunoglobulin heavy chain constant region most preferably comprises an immunoglobulin hinge region, a CH2 domain and a CH3 domain. The preferred Fc domain according to this invention consists thus of the hinge-CH2-CH3 domain.

As used herein, the term immunoglobulin “hinge region” is understood to mean an entire immunoglobulin hinge region or at least a portion of the immunoglobulin hinge region sufficient to form one or more disulfide bonds with a second immunoglobulin hinge region.

As used herein, the term "signal sequence" is understood to mean a segment which directs the secretion of the fusion protein and thereafter is cleaved following translation in the host cell. The signal sequence of the invention is a polynucleotide which encodes an amino acid

sequence which initiates transport of a protein across the membrane of the endoplasmic reticulum. Signal sequences which are useful in the invention include antibody light chain signal sequences, e.g., antibody 14.18 (Gillies *et. al.* (1989) J. Immunol. Meth. 125 : 191), and any other signal sequences which are known in the art (see, e.g., Watson, 1984, Nucleic Acid Research 12:5145).

The term "mutant or variant" used with respect to a particular protein encompasses any molecule such as a truncated or other derivative of the relevant protein which retains substantially the same activity in humans as the relevant protein. Such other derivatives can be prepared by the addition, deletion, substitution, or rearrangement of amino acids or by chemical modifications thereof.

It is contemplated that suitable immunoglobulin heavy chain constant regions may be derived from antibodies belonging to each of the immunoglobulin classes referred to as IgA, IgD, IgE, IgG, and IgM, however, immunoglobulin heavy chain constant regions from the IgG class are preferred.

Furthermore, it is contemplated that immunoglobulin heavy chain constant regions may be derived from any of the IgG antibody subclasses referred to in the art as IgG1, IgG2, IgG3, and IgG4. Immunoglobulin heavy chain constant region domains have cross-homology among the immunoglobulin classes. For example, the CH2 domain of IgG is homologous to the CH2 domain of IgA and IgD, and to the CH3 domain of IgM and IgE. The choice of appropriate immunoglobulin heavy chain constant regions is discussed in detail in US 5,541,087 and 5,726,044. The choice of particular immunoglobulin heavy chain constant region sequences from certain immunoglobulin classes and subclasses to achieve a particular result is considered to be within the level of skill in the art. It may be useful, in some circumstances, to modify the immunoglobulin heavy chain constant region, for example, by mutation, deletion or other changes mediated by genetic engineering or other approaches, so that certain activities, such as complement fixation or stimulation of antibody-dependent cell-mediated cytotoxicity (ADCC) are reduced or eliminated.

The Fc region is considered non- or weakly immunogenic if the immunoglobulin heavy chain constant region fails to generate a detectable antibody response.

Furthermore, it is contemplated that substitution or deletion of amino acids within the immunoglobulin heavy chain constant regions may be useful in the practice of the

invention. One example may include introducing amino acid substitutions in the upper CH2 region to create a Fc variant with reduced affinity for Fc receptors (Cole *et al.* (1997) J. Immunol.159:3613). An antibody-based fusion protein with an enhanced *in vivo* circulating half-life can be obtained by constructing a fusion protein having reduced binding affinity for a Fc receptor, and avoiding the use of sequences from antibody isotypes that bind to Fc receptors (WO 99/43713). For example, of the four known IgG isotypes, IgG1 (Cγ1) and IgG3 (Cγ3) are known to bind FcRγ1 with high affinity, whereas IgG4 has a 10-fold lower binding affinity, and IgG2 (Cγ2) does not bind to FcRγ1. Thus, an antibody-based fusion protein with reduced binding affinity for a Fc receptor could be obtained by constructing a fusion protein with a Cγ2 constant region (Fc region) or a Cγ4 Fc region, and avoiding constructs with a Cγ1 Fc region or a Cγ3 Fc region. An antibody-based fusion protein with an enhanced *in vivo* circulating half-life can be obtained by modifying sequences necessary for binding to Fc receptors in isotypes that have binding affinity for an Fc receptor, in order to reduce or eliminate binding.

The important sequences for FcγR binding are Leu-Leu-Gly-Gly (residues 234 through 237 in Cγ1), located in the CH2 domain adjacent to the hinge (Canfield and Morrison, J. Exp. Med. 173: 1483-1491 (1991)). Another important structural component necessary for effective FcR binding is the presence of an N-linked carbohydrate chain covalently bound to Asn_{297}. Enzymatic removal of this structure or mutation of the Asn residue effectively abolish, or at least dramatically reduce, binding to all classes of FcγR.

The resulting antibody-based fusion proteins have a longer *in vivo* circulating half-life than the unlinked second non-immunoglobulin protein. Dimerization of a ligand can increase the apparent binding affinity between the ligand and its receptor. For instance, if one X moiety of an Fc-X fusion protein can bind to a receptor on a cell with a certain affinity, the second X moiety of the same Fc-Interferon-alpha fusion protein may bind to a second receptor on the same cell with a much higher avidity (apparent affinity). This may occur because of the physical proximity of the second X moiety to the receptor after the first X moiety already is bound. In the case of an antibody binding to an antigen, the apparent affinity may be increased by at least ten thousand-fold. Each protein subunit, i.e., "X," has its own independent function so that in a multivalent molecule, the functions of the protein subunits

may be additive or synergistic. Thus, fusion of the normally dimeric Fc molecule or another antibody fragment to a polypeptide X may increase the activity of X.

Nucleic acid sequences encoding, and amino acid sequences defining a human immunoglobulin Fc region, especially a Fcγ1, Fcγ2 and Fcγ3, useful in the practice of the invention are set forth in in the prior, such as disclosed in (WO 00/40615, WO 00/69913, WO 00/24782) or in the Genbank and/or EMBL databases, for example, AF045536.1 *(Macaca fuscicularis),* AF045537.1 *(Macaca mulatta),* AB016710 *(Felix catus),* K00752 *(Oryctolagus cuniculus),* U03780 *(Sus scrofa), 248947 (Camelus dromedarius),* X62916 *(Bos taurus),* L07789 *(Mustela vison), X69797 (Ovis aries),* U17 *166 (Cricetulus migratorius),* X07189 *(Rattus rattus),* AF57619.1 *(Trichosurus vulpecula),* or AF035195 *(Monodelphis domestica).*

Thus, vectors reported earlier (Lo *et al.* (1998) Protein Engeneering 11:495-500) were modified by replacing the human IgGl Fc sequence with sequences from cDNA encoding the mouse IgG2a Fc (US 5,726,044).

The invention encompasses mutations in the immunoglobulin component which eliminate undesirable properties of the native immunoglobulin, such as Fc receptor binding and/or introduce desirable properties such as stability. For example, Angal S., King D.J., Bodmer M.W., Turner A., Lawson A.D.G., Roberts G., Pedley B. and Adair R., Molecular Immunology,130, pp105-108, 1993, describe an IgG4 molecule where residue 241 (Kabat numbering) is altered from serine to proline. This change increases the serum half-life of the IgG4 molecule. Canfield S.M. and Morrison S.L., Journal of Experimental Medicine vol173pp1483-149 1, describe the alteration of residue 248 (Kabat numbering) from leucine to glutamate in IgG3 and from glutamate to leucine in mouse IgG2b. Substitution of leucine for glutamate in the former decreases the affinity of the immunoglobulin molecule concerned for the FcγR1 receptor, and substitution of glutamate for leucine in the latter increases the affinity. EP 0307 434 discloses various mutations including an L to E mutation at residue 248 (Kabat numbering) in IgG. The constant domain(s) or fragment thereof is preferably the whole or a substantial part of the constant region of the heavy chain of human IgG. The IgG component suitably comprises the CH2 and CH3 domains and the hinge region including cysteine residues contributing to inter-heavy chain disulphide bonding. For example

when the IgG component is derived from IgG4 it includes cysteine residues 8 and 11 of the IgG4 hinge region (Pinck J.R. and Milstein C., Nature, 121, 6pp941-942, 1967).

The process of the invention may be performed by conventional recombinant techniques such as described in Maniatis *et. al.* (Molecular Cloning - A Laboratory Manual; Cold Spring Harbor, 1982) and DNA Cloning Vols I, II and III (D.M. Glover ed., IRL Press Ltd) or Sambrook et al. (1989, Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory Press, NY, USA).

In particular, the process may comprise the steps of:
(i) preparing a replicable expression vector capable, in a host cell, of expressing a DNA polymer comprising a nucleotide sequence that encodes said compound;
(ii) transforming a host cell with said vector;
(iii) culturing said transformed host cell under conditions permitting expression of said DNA polymer to produce said compound; and
(iv) recovering said compound.

The invention also provides a process for preparing the DNA polymer by the condensation of appropriate mono-, di- or oligomeric nucleotide units. The preparation may be carried out chemically, enzymatically, or by a combination of the two methods, *in virro* or *in vivo* as appropriate. Thus, the DNA polymer may be prepared by the enzymatic ligation of appropriate DNA fragments, by conventional methods such as those described by D. M. Roberts et al in Biochemistry 1985, 24, 5090-5098. The DNA fragments may be obtained by digestion of DNA containing the required sequences of nucleotides with appropriate restriction enzymes, by chemical synthesis, by enzymatic polymerisation on DNA or RNA templates, or by a combination of these methods. Digestion with restriction enzymes may be performed in an appropriate buffer at a temperature of 20°-70°C with 0.1- 10μg DNA. Enzymatic polymerisation of DNA may be carried out *in vitro* using a DNA polymerase such as DNA polymerase I (Klenow fragment) in an appropriate buffer containing the nucleoside uiphosphates dATP, dCTP, dGTP and dlTP as required at a temperature of 10°-37°C, generally in a volume of 5Oμl or less. Enzymatic ligation of DNA fragments may be carried out using a DNA ligase such as T4 DNA ligase in an appropriate buffer at a temperature of 40°C to ambient, generally in a volume of 50μl or less.

The chemical synthesis of the DNA polymer or fragments may be carried out by conventional phosphotriester, phosphite or phosphoramidite chemistry, using solid phase techniques such as those described in "Chemical and Enzymatic Synthesis of Gene Fragments - A Laboratory Manual" (ed. H.G. Gassen and A. Lang), Verlag Chemie, Weinheim (1982),or in other scientific publications, for example M.J. Gait, H.W.D. Matthes, M. Singh, B.S. Sproat, and R.C. Titmas, Nucleic Acids Research, 1982,10, 6243; B.S. Sproat and W. Bannwarth, Tetrahedron Letters, 1983,24,5771; M.D. Matteucci and M.H Caruthers, Tetrahedron Letters, 1980,21,719; M.D. Matteucci and M.H. Caruthers, Journal of the American Chemical Society, 1981,103, 3185; S.P. Adams et *al.*, Journal of the American Chemical Society,1983, 105, 661; N.D. Sinha, J. Biemat, J. McMannus, and H. Koester, Nucleic Acids Research, 1984, 12,4539; and H.W.D. Matthes et al., EMBO Journal, 1984,3,801. Preferably an automated DNA synthesizer is employed.

The DNA molecules may be obtained by the digestion with suitable restriction enzymes of vectors carrying the required coding sequences or by use of polymerase chain reaction technology. The precise structure of the DNA molecules and the way in which they are obtained depends upon the structure of the desired product. The design of a suitable strategy for the construction of the DNA molecule coding for the compound is a routine matter for the skilled worker in the art.

The expression of the DNA polymer encoding the compound in a recombinant host cell may be carried out by means of a replicable expression vector capable, in the host cell, of expressing the DNA polymer. The expression vector is novel and also forms part of the invention. The replicable expression vector may be prepared in accordance with the invention, by cleaving a vector compatible with the host cell to provide a linear DNA segment having an intact replicon, and combining said linear segment with one or more DNA molecules which, together with said linear segment, encode the compound, under ligating conditions. The ligation of the linear segment and more than one DNA molecule may be carried out simultaneously or sequentially as desired. Thus, the DNA polymer may be preformed or formed during the construction of the vector, as desired. A useful expression vector is described at Lo *et al.* (1988) Protein Engineeering 11:495, in which the transcription of the Fc-X gene utilizes the enhancer/promoter of the human cytomegalovirus and the SV40

polyadenylation signal. Suitable vectors include plasmids, bacteriophages, cosmids and recombinant viruses derived from, for example, baculoviruses, vaccinia or Semliki Forest virus. Thus, vectors reported earlier (Lo *et al.* (1998) Protein Engeneering 11:495-500) were modified by replacing the human IgG1 Fc sequence with sequences from cDNA encoding the mouse IgG2a Fc (US 5,726,044).

The choice of vector will be determined in part by the host cell, which may be prokaryotic, such as *E. coli,* or eukaryotic, such as mouse C127, mouse myeloma, Chinese hamster ovary, COS or Hela cells, fungi e.g. filamentous fungi or unicellular yeast or an insect cell such as Drosophila. Currently preferred host cells for use in the invention include immortal hybridoma cells, NS/0 myeloma cells, 293 cells, Chinese hamster ovary cells, HELA cells, and COS cells. The host cell may also be a transgenic animal.

Polymerisation and ligation may be performed as described above for the preparation of the DNA polymer. Digestion with restriction enzymes may be performed in an appropriate buffer at a temperature of 20°- 70°C with 0.1 -10μg DNA. The recombinant host cell is prepared, in accordance with the invention, by transforming a host cell with a replicable expression vector of the invention under transforming conditions. Suitable transforming conditions are conventional and are described in, for example, Maniatis et al., cited above, or "DNA Cloning" Vol. II, D.M. Glover ed., IRL Press Ltd, 1985. The choice of transforming conditions is determined by the host cell. Thus, a bacterial host such as *E. coli* may be treated with a solution of $CaCl_2$ (Cohen er al, Proc. Nat. Acad. Sci., 1973,69,2110) or with a solution comprising a mixture of RbCl, $MnCl_2$, potassium acetate and glycerol, and then with 3-[N-morpholinol-propane- sulphonic acid, RbCl and glycerol. Mammalian cells in culture may be transformed by calcium co-precipitation of the vector DNA onto the cells. The invention also extends to a host cell transformed or transfected with a replicable expression vector of the invention. Culturing the transformed host cell under conditions permitting expression of the DNA polymer is carried out conventionally, as described in, for example, Maniatis et al. and "DNA Cloning" cited above. Thus, preferably the cell is supplied with nutrient and cultured at a temperature below 45°C. The expression product is recovered by conventional methods according to the host cell. Thus, where the host cell is bacterial, such as *E. coli* it may be lysed physically, chemically or enzymatically and the protein product isolated from the resulting lysate. If the product is to be secreted from the bacterial cell it may

be recovered from the periplasmic space or the nutrient medium. Where the host cell is mammalian, the product may generally be isolated from the nutrient medium. The DNA polymer may be assembled into vectors designed for isolation of stable transformed mammalian cell lines expressing the product: e.g. bovine papillomavirus vectors or amplified vectors in Chinese hamster ovary cells (DNA cloning Vol.11 D.M. Glover ed. IRL Press 1985; Kaufman, R.J., Molecular and Cellular Biology 5, 1750-1759, 1985; Pavlakis G.N. and Hamer, D.H., Proceedings of the National Academy of Sciences (USA) 80,397-401,1983; Goeddel, D.V. et al., and EP 0 093 619,1983).

The immunoconjugates of the invention may comprise linker molecules. The linker is preferably made up of amino acids linked together by peptide bonds. Thus, in preferred embodiments, the linker is made up of from 1 to 20 amino acids linked by peptide bonds, wherein the amino acids are selected from the 20 naturally occurring amino acids. Some of these amino acids may be glycosylated, as is well understood by those in the art. In a more preferred embodiment, the 1 to 20 amino acids are selected from glycine, alanine, proline, asparagine, glutamine, and lysine. Even more preferably, a linker is made up of a majority of amino acids that are sterically unhindered, such as glycine and alanine. Thus, preferred linkers are polyglycines such as

polyGly (particularly $(Gly)_2$ - $(Gly)_7$),

poly(Gly-Ala),

polyAla .

Other specific examples of suitable linkers are:

$(Gly)_3Lys(Gly)_4$

$(Gly)_3AsnGlySer(Gly)_2$

$(Gly)_3Cys(Gly)_4$ and

GlyProAsnGlyGly

Combinations of Gly and Ala are also preferred. The linkers shown here are exemplary; linkers within the scope of this invention may be much longer and may include other residues. Non-peptide linkers are also possible. The peptide linkers may be altered to form derivatives in the same manner as described above.

Preferred linkers of the invention are not or less immunogenic. Most of the above-cited linker peptides are at least less immunogenic. However it is possible that creating the linkage between an antibody or a sFv, Fab, Fab' or F(ab')2 or a Fc domain and the target protein via a

linker peptide molecule as mentioned above, new immunogenic epitopes may be newly created within the linkage region resulting in an immunoconjugate which has an increased immunogenicity compared to the immunogenicity of the single (de-immunized) components. This situation can also be extended to fusion protein having no linker molecule. Therefore, the invention also relates to de-immunized regions of a fusion protein according to the invention, the so-called fusion or junction regions. When fusing a first protein molecule with a second protein molecule (which also can be a linker molecule) via the C- and N-terminals a sequence region is created that is artificial and, thus, was usually not yet seen by the immune system. This region is deemed is be immunogenic. The region of amino acid residues comprise according to the invention approximately 10 residues of each protein terminal (N- or C-terminal). The complete fusion region comprises, therefore, about 20 amino acid residues, preferably 2 – 16, more preferably 2 – 10 (which is 1 – 8 and 1 – 5 amino acid residues, respectively, of each fusion partner).

The invention includes also further Fc variants. Such further Fc variants, one may remove one or more sites of a native Fc that provide structural features or functional activity not required by the fusion molecules of this invention. One may remove these sites by, for example, substituting or deleting residues, inserting residues into the site, or truncating portions containing the site. The inserted or substituted residues may also be altered amino acids, such as peptidomimetics or D- amino acids. For example, one or more glycosylation sites may be removed. Residues that are typically glycosylated (e.g., asparagine) may confer cytolytic response. Such residues may be deleted or substituted with unglycosylated residues (e.g., alanine). ADCC site as well as sites involved in interaction with complement, such as the Clq binding site, may also be removed if there is a specific need.

The invention includes also derivatives of the target polypeptide (X) of the invention. Such derivatives may further improve the solubility, absorption, biological half life, and the like of (X). The modified (X) may alternatively eliminate or attenuate any undesirable side-effect and the like. Exemplary derivatives include also compounds in which (X) or some portion thereof is cyclic. For example, the peptide portion may be modified to contain two or more Cys residues (e.g., in the linker), which could cyclize by disulfide bond formation. The compound is cross-linked or is rendered capable of cross-linking between molecules. For example, the

peptide portion may be modified to contain one Cys residue and thereby be able to form an intermolecular disulfide bond with a like molecule.

In a final aspect the present invention relates to pharmaceutical compositions comprising said biologically active proteins obtainable by the methods disclosed in the present invention, and methods for therapeutic treatment of humans using the modified molecules and pharmaceutical compositions.

Therapeutic compositions of the present invention contain a physiologically tolerable carrier together with the relevant agent as described herein, dissolved or dispersed therein as an active ingredient. As used herein. the terms "pharmaceutically acceptable", "physiologically tolerable" and grammatical variations thereof, as they refer to compositions, carriers, diluents and reagents, are used interchangeably and represent that the materials are capable of administration to or upon a mammal without the production of undesirable physiological effects such as nausea, dizziness, gastric upset and the like. The preparation of a pharmacological composition that contains active ingredients dissolved or dispersed therein is well understood in the art and need not be limited based on formulation. Typically, such compositions are prepared as injectables either as liquid solutions or suspensions, however, solid forms suitable for solution, or suspensions, in liquid prior to use can also be prepared. The preparation can also be emulsified. The active ingredient can be mixed with excipients which are pharmaceutically acceptable and compatible with the active ingredient and in amounts suitable for use in the therapeutic methods described herein. Suitable excipients are, for example, water, saline, dextrose, glycerol, ethanol or the like and combinations thereof. In addition, if desired, the composition can contain minor amounts of auxiliary substances such as wetting or emulsifying agents, pH buffering agents and the like which enhance the effectiveness of the active ingredient.

The therapeutic composition of the present invention can include pharmaceutically acceptable salts of the components therein. Pharmaceutically acceptable salts include the acid addition salts (formed with the free amino groups of the polypeptide) that are formed with inorganic acids such as. for example, hydrochloric or phosphoric acids, or such organic acids as acetic, tartaric, mandelic and the like. Salts formed with the free carboxyl groups can also be derived from inorganic bases such as, for example, sodium, potassium, ammonium, calcium or ferric hydroxides, and such organic bases as isopropylamine, trimethylamine, 2-ethylamino ethanol, histidine, procaine and the like. Particularly preferred is the HCI salt when used in the

preparation of cyclic polypeptide αv antagonists. Physiologically tolerable carriers are well known in the art. Exemplary of liquid carriers are sterile aqueous solutions that contain no materials in addition to the active ingredients and water, or contain a buffer such as sodium phosphate at physiological pH value, physiological saline or both, such as phosphate-buffered saline. Still further, aqueous carriers can contain more than one buffer salt, as well as salts such as sodium and potassium chlorides, dextrose, polyethylene glycol and other solutes. Liquid compositions can also contain liquid phases in addition to and to the exclusion of water. Exemplary of such additional liquid phases are glycerin. vegetable oils such as cottonseed oil, and water-oil emulsions.

Typically, a therapeutically effective amount of a modified immunoglobulin in the form of an modified antibody or antibody fragment according to the invention is an amount such that when administered in physiologically tolerable composition is sufficient to achieve a plasma concentration of from about 0.01 microgram (μg) per milliliter (ml) to about 100 μg/ml, preferably from about 1 μg/ml to about 5 μg/ml and usually about 5 μg/ml. Stated differently. the dosage can vary from about 0.1 mg/kg to about 300 mg/kg, preferably from about 0.2 mg/kg to about 200 mg/kg, most preferably from about 0.5 mg/kg to about 20 mg/kg, in one or more dose administrations daily for one or several days. Where the immunotherapeutic agent is in the form of a fragment of a monoclonal antibody or a conjugate, the amount can readily be adjusted based on the mass of the fragment / conjugate relative to the mass of the whole antibody. A preferred plasma concentration in molarity is from about 2 micromolar (μM) to about 5 millimolar (mM) and preferably, about 100 μM to 1 mM antibody antagonist.

A therapeutically effective amount of an agent according of this invention which is a non-immunotherapeutic peptide or a protein is typically an amount of such a molecule such that when administered in a physiologically tolerable composition is sufficient to achieve a plasma concentration of from about 0.1 microgram (μg) per milliliter (ml) to about 200 μg/ml, preferably from about 1 μg/ml to about 150 μg/ml. Based on a protein having a mass of about 500 grams per mole, the preferred plasma concentration in molarity is from about 2 micromolar (μM) to about 5 millimolar (mM) and preferably about 100 μM to 1 mM polypeptide antagonist.

The pharmaceutical compositions of the invention can comprise phrase encompasses treatment of a subject with agents that reduce or avoid side effects associated with the combination

therapy of the present invention ("adjunctive therapy"), including, but not limited to, those agents, for example, that reduce the toxic effect of anticancer drugs. Said adjunctive agents prevent or reduce the incidence of nausea and vomiting associated with chemotherapy, radiotherapy or operation, or reduce the incidence of infection associated with the administration of myelosuppressive anticancer drugs. Adjunctive agents are well known in the art. The modified proteins according to the invention can additionally administered with adjuvants like BCG and immune system stimulators.

Furthermore, the compositions may include immunotherapeutic agents, chemotherapeutic agents and anti-neoplastic agents which may contain cytotoxic effective radio labeled isotopes, or other cytotoxic agents, such as a cytotoxic peptides (e.g. cytokines) or cytotoxic drugs and the like. The typical dosage of an active agent, which is a preferably a chemical antagonist or a (chemical) chemotherapeutic agent according to the invention (neither an immunotherapeutic agent nor a non-immunotherapeutic peptide/protein) is 10 mg to 1000 mg, preferably about 20 to 200 mg, and more preferably 50 to 100 mg per kilogram body weight per day.

The following examples describe the in invention in more detail. However, this listing does not limit the invention.

EXAMPLE 1:

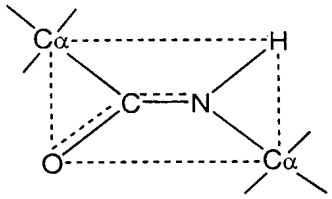
The following example describes in detail a preferred method for identification of immunogenic sequence regions (T-cell epitopes) within the sequences of the fusion proteins as disclosed in this invention. However, it should be pointed out, that said molecules can be obtained by other known methods.

The identification of T-cell epitopes of the molecules which were modified in order to obtain the immunoconjugates according to the present invention can be achieved by different methods which are described in the prior art (WO 92/10755 and WO 96/40792 (Novo Nordisk), EP 0519 596 (Merck & Co.), EP 0699 755(Centro de Immunologia Moelcular), WO 98/52976 and WO 98/59244 (Biovation Ltd.) or related methods.

Advantageous immunoconjugates, however, can be obtained if the identification of said epitopes is realized by the following new method which is described herewith in detail:

There are a number of factors that play important roles in determining the total structure of a protein, polypeptide or immunoglobulin. First, the peptide bond, i.e., that bond which joins the amino acids in the chain together, is a covalent bond. This bond is planar in structure, essentially a substituted amide. An "amide" is any of a group of organic compounds containing the grouping -CONH-.

The planar peptide bond linking Cα of adjacent amino acids may be represented as depicted below:



Because the O=C and the C-N atoms lie in a relatively rigid plane, free rotation does not occur about these axes. Hence, a plane schematically depicted by the interrupted line is sometimes referred to as an "amide" or "peptide plane" plane wherein lie the oxygen (O), carbon (C), nitrogen (N), and hydrogen (H) atoms of the peptide backbone. At opposite corners of this amide plane are located the Cα atoms. Since there is substantially no rotation about the O=C and C-N atoms in the peptide or amide plane, a polypeptide chain thus comprises a series of planar peptide linkages joining the Cα atoms.

A second factor that plays an important role in defining the total structure or conformation of a polypeptide or protein is the angle of rotation of each amide plane about the common Cα linkage. The terms "angle of rotation" and "torsion angle" are hereinafter regarded as equivalent terms. Assuming that the O, C, N, and H atoms remain in the amide plane (which is usually a valid assumption, although there may be some slight deviations from planarity of these atoms for some conformations), these angles of rotation define the N and R polypeptide's backbone conformation, i.e., the structure as it exists between adjacent residues. These two angles are known as ϕ and ψ. A set of the angles ϕ_1, ψ_1, where the subscript i represents a particular residue of a polypeptide chain, thus effectively defines the polypeptide secondary structure. The conventions used in defining the ϕ, ψ angles, i.e., the reference points at which the amide planes form a zero degree angle, and the definition of which angle is ϕ, and which angle is ψ, for a given polypeptide, are defined in the literature. See, e.g.,

Ramachandran et al. *Adv. Prot. Chem.* 23:283-437 (1968), at pages 285-94, which pages are incorporated herein by reference.

The present method can be applied to any protein, and is based in part upon the discovery that in humans the primary Pocket 1 anchor position of MHC Class II molecule binding grooves has a well designed specificity for particular amino acid side chains. The specificity of this pocket is determined by the identity of the amino acid at position 86 of the beta chain of the MHC Class II molecule. This site is located at the bottom of Pocket 1 and determines the size of the side chain that can be accommodated by this pocket. Marshall, K.W., *J. Immunol.*, 152:4946-4956 (1994). If this residue is a glycine, then all hydrophobic aliphatic and aromatic amino acids (hydrophobic aliphatics being: valine, leucine, isoleucine, methionine and aromatics being: phenylalanine, tyrosine and tryptophan) can be accommodated in the pocket, a preference being for the aromatic side chains. If this pocket residue is a valine, then the side chain of this amino acid protrudes into the pocket and restricts the size of peptide side chains that can be accommodated such that only hydrophobic aliphatic side chains can be accommodated. Therefore, in an amino acid residue sequence, wherever an amino acid with a hydrophobic aliphatic or aromatic side chain is found, there is the potential for a MHC Class II restricted T-cell epitope to be present. If the side-chain is hydrophobic aliphatic, however, it is approximately twice as likely to be associated with a T-cell epitope than an aromatic side chain (assuming an approximately even distribution of Pocket 1 types throughout the global population).

A computational method embodying the present invention profiles the likelihood of peptide regions to contain T-cell epitopes as follows:

(1) The primary sequence of a peptide segment of predetermined length is scanned, and all hydrophobic aliphatic and aromatic side chains present are identified. (2)The hydrophobic aliphatic side chains are assigned a value greater than that for the aromatic side chains; preferably about twice the value assigned to the aromatic side chains, e.g., a value of 2 for a hydrophobic aliphatic side chain and a value of 1 for an aromatic side chain. (3) The values determined to be present are summed for each overlapping amino acid residue segment (window) of predetermined uniform length within the peptide, and the total value for a particular segment (window) is assigned to a single amino acid residue at an intermediate position of the segment (window), preferably to a residue at about the midpoint of the sampled segment (window). This procedure is repeated for each sampled overlapping amino acid residue segment (window). Thus, each amino acid residue of the peptide is assigned a value

that relates to the likelihood of a T-cell epitope being present in that particular segment (window). (4) The values calculated and assigned as described in Step 3, above, can be plotted against the amino acid coordinates of the entire amino acid residue sequence being assessed. (5) All portions of the sequence which have a score of a predetermined value, e.g., a value of 1, are deemed likely to contain a T-cell epitope and can be modified, if desired.

This particular aspect of the present invention provides a general method by which the regions of peptides likely to contain T-cell epitopes can be described. Modifications to the peptide in these regions have the potential to modify the MHC Class II binding characteristics.

According to another aspect of the present invention, T-cell epitopes can be predicted with greater accuracy by the use of a more sophisticated computational method which takes into account the interactions of peptides with models of MHC Class II alleles.

The computational prediction of T-cell epitopes present within a peptide according to this particular aspect contemplates the construction of models of at least 42 MHC Class II alleles based upon the structures of all known MHC Class II molecules and a method for the use of these models in the computational identification of T-cell epitopes, the construction of libraries of peptide backbones for each model in order to allow for the known variability in relative peptide backbone alpha carbon (Cα) positions, the construction of libraries of amino-acid side chain conformations for each backbone dock with each model for each of the 20 amino-acid alternatives at positions critical for the interaction between peptide and MHC Class II molecule, and the use of these libraries of backbones and side-chain conformations in conjunction with a scoring function to select the optimum backbone and side-chain conformation for a particular peptide docked with a particular MHC Class II molecule and the derivation of a binding score from this interaction.

Models of MHC Class II molecules can be derived via homology modeling from a number of similar structures found in the Brookhaven Protein Data Bank ("PDB"). These may be made by the use of semi-automatic homology modeling software (Modeller, Sali A. & Blundell TL., 1993. *J. Mol Biol* 234:779-815) which incorporates a simulated annealing function, in conjunction with the CHARMm force-field for energy minimisation (available from Molecular Simulations Inc., San Diego, Ca.). Alternative modeling methods can be utilized as well.

The present method differs significantly from other computational methods which use libraries of experimentally derived binding data of each amino-acid alternative at each position in the binding groove for a small set of MHC Class II molecules (Marshall, K.W., *et al.*, *Biomed.*

Pept. Proteins Nucleic Acids, 1(3):157-162) (1995) or yet other computational methods which use similar experimental binding data in order to define the binding characteristics of particular types of binding pockets within the groove, again using a relatively small subset of MHC Class II molecules, and then 'mixing and matching' pocket types from this pocket library to artificially create further 'virtual' MHC Class II molecules (Sturniolo T., et al., *Nat. Biotech*, 17(6): 555-561 (1999). Both prior methods suffer the major disadvantage that, due to the complexity of the assays and the need to synthesize large numbers of peptide variants, only a small number of MHC Class II molecules can be experimentally scanned. Therefore the first prior method can only make predictions for a small number of MHC Class II molecules. The second prior method also makes the assumption that a pocket lined with similar amino-acids in one molecule will have the same binding characteristics when in the context of a different Class II allele and suffers further disadvantages in that only those MHC Class II molecules can be 'virtually' created which contain pockets contained within the pocket library. Using the modeling approach described herein, the structure of any number and type of MHC Class II molecules can be deduced, therefore alleles can be specifically selected to be representative of the global population. In addition, the number of MHC Class II molecules scanned can be increased by making further models further than having to generate additional data via complex experimentation.

The use of a backbone library allows for variation in the positions of the Cα atoms of the various peptides being scanned when docked with particular MHC Class II molecules. This is again in contrast to the alternative prior computational methods described above which rely on the use of simplified peptide backbones for scanning amino-acid binding in particular pockets. These simplified backbones are not likely to be representative of backbone conformations found in 'real' peptides leading to inaccuracies in prediction of peptide binding. The present backbone library is created by superposing the backbones of all peptides bound to MHC Class II molecules found within the Protein Data Bank and noting the root mean square (RMS) deviation between the Cα atoms of each of the eleven amino-acids located within the binding groove. While this library can be derived from a small number of suitable available mouse and human structures (currently 13), in order to allow for the possibility of even greater variability, the RMS figure for each C"-α position is increased by 50%. The average Cα position of each amino-acid is then determined and a sphere drawn around this point whose radius equals the RMS deviation at that position plus 50%. This sphere represents all allowed Cα positions.

Working from the Cα with the least RMS deviation (that of the amino-acid in Pocket 1 as mentioned above, equivalent to Position 2 of the 11 residues in the binding groove), the sphere is three-dimensionally gridded, and each vertex within the grid is then used as a possible location for a Cα of that amino-acid. The subsequent amide plane, corresponding to the peptide bond to the subsequent amino-acid is grafted onto each of these Cαs and the ϕ and ψ angles are rotated step-wise at set intervals in order to position the subsequent Cα. If the subsequent Cα falls within the 'sphere of allowed positions' for this Cα than the orientation of the dipeptide is accepted, whereas if it falls outside the sphere then the dipeptide is rejected. This process is then repeated for each of the subsequent Cα positions, such that the peptide grows from the Pocket 1 Cα 'seed', until all nine subsequent Cαs have been positioned from all possible permutations of the preceding Cαs. The process is then repeated once more for the single Cα preceding pocket 1 to create a library of backbone Cα positions located within the binding groove.

The number of backbones generated is dependent upon several factors: The size of the 'spheres of allowed positions'; the fineness of the gridding of the 'primary sphere' at the Pocket 1 position; the fineness of the step-wise rotation of the ϕ and ψ angles used to position subsequent Cαs. Using this process, a large library of backbones can be created. The larger the backbone library, the more likely it will be that the optimum fit will be found for a particular peptide within the binding groove of an MHC Class II molecule. Inasmuch as all backbones will not be suitable for docking with all the models of MHC Class II molecules due to clashes with amino-acids of the binding domains, for each allele a subset of the library is created comprising backbones which can be accommodated by that allele. The use of the backbone library, in conjunction with the models of MHC Class II molecules creates an exhaustive database consisting of allowed side chain conformations for each amino-acid in each position of the binding groove for each MHC Class II molecule docked with each allowed backbone. This data set is generated using a simple steric overlap function where a MHC Class II molecule is docked with a backbone and an amino-acid side chain is grafted onto the backbone at the desired position. Each of the rotatable bonds of the side chain is rotated step-wise at set intervals and the resultant positions of the atoms dependent upon that bond noted. The interaction of the atom with atoms of side-chains of the binding groove is noted and positions are either accepted or rejected according to the following criteria: The sum total of the overlap of all atoms so far positioned must not exceed a pre-determined value.

Thus the stringency of the conformational search is a function of the interval used in the step-wise rotation of the bond and the pre-determined limit for the total overlap. This latter value can be small if it is known that a particular pocket is rigid, however the stringency can be relaxed if the positions of pocket side-chains are known to be relatively flexible. Thus allowances can be made to imitate variations in flexibility within pockets of the binding groove. This conformational search is then repeated for every amino-acid at every position of each backbone when docked with each of the MHC Class II molecules to create the exhaustive database of side-chain conformations.

A suitable mathematical expression is used to estimate the energy of binding between models of MHC Class II molecules in conjunction with peptide ligand conformations which have to be empirically derived by scanning the large database of backbone/side-chain conformations described above. Thus a protein is scanned for potential T-cell epitopes by subjecting each possible peptide of length varying between 9 and 20 amino-acids (although the length is kept constant for each scan) to the following computations: An MHC Class II molecule is selected together with a peptide backbone allowed for that molecule and the side-chains corresponding to the desired peptide sequence are grafted on. Atom identity and interatomic distance data relating to a particular side-chain at a particular position on the backbone are collected for each allowed conformation of that amino-acid (obtained from the database described above). This is repeated for each side-chain along the backbone and peptide scores derived using a scoring function. The best score for that backbone is retained and the process repeated for each allowed backbone for the selected model. The scores from all allowed backbones are compared and the highest score is deemed to be the peptide score for the desired peptide in that MHC Class II model. This process is then repeated for each model with every possible peptide derived from the protein being scanned, and the scores for peptides versus models are displayed.

In the context of the present invention, each ligand presented for the binding affinity calculation is an amino-acid segment selected from a peptide or protein as discussed above. Thus, the ligand is a selected stretch of amino acids about 9 to 20 amino acids in length derived from a peptide, polypeptide or protein of known sequence. The terms "amino acids" and "residues" are hereinafter regarded as equivalent terms. The ligand, in the form of the consecutive amino acids of the peptide to be examined grafted onto a backbone from the backbone library, is positioned in the binding cleft of an MHC Class II molecule from the MHC Class II molecule model library via the coordinates of the C"-α atoms of the peptide

backbone and an allowed conformation for each side-chain is selected from the database of allowed conformations. The relevant atom identities and interatomic distances are also retrieved from this database and used to calculate the peptide binding score. Ligands with a high binding affinity for the MHC Class II binding pocket are flagged as candidates for site-directed mutagenesis. Amino-acid substitutions are made in the flagged ligand (and hence in the protein of interest) which is then retested using the scoring function in order to determine changes which reduce the binding affinity below a predetermined threshold value. These changes can then be incorporated into the protein of interest to remove T-cell epitopes.

Binding between the peptide ligand and the binding groove of MHC Class II molecules involves non-covalent interactions including, but not limited to: hydrogen bonds, electrostatic interactions, hydrophobic (lipophilic) interactions and Van der Walls interactions. These are included in the peptide scoring function as described in detail below. It should be understood that a hydrogen bond is a non-covalent bond which can be formed between polar or charged groups and consists of a hydrogen atom shared by two other atoms. The hydrogen of the hydrogen donor has a positive charge where the hydrogen acceptor has a partial negative charge. For the purposes of peptide/protein interactions, hydrogen bond donors may be either nitrogens with hydrogen attached or hydrogens attached to oxygen or nitrogen. Hydrogen bond acceptor atoms may be oxygens not attached to hydrogen, nitrogens with no hydrogens attached and one or two connections, or sulphurs with only one connection. Certain atoms, such as oxygens attached to hydrogens or imine nitrogens (e.g. C=NH) may be both hydrogen acceptors or donors. Hydrogen bond energies range from 3 to 7 Kcal/mol and are much stronger than Van der Waal's bonds, but weaker than covalent bonds. Hydrogen bonds are also highly directional and are at their strongest when the donor atom, hydrogen atom and acceptor atom are co-linear. Electrostatic bonds are formed between oppositely charged ion pairs and the strength of the interaction is inversely proportional to the square of the distance between the atoms according to Coulomb's law. The optimal distance between ion pairs is about 2.8Å. In protein/peptide interactions, electrostatic bonds may be formed between arginine, histidine or lysine and aspartate or glutamate. The strength of the bond will depend upon the pKa of the ionizing group and the dielectric constant of the medium although they are approximately similar in strength to hydrogen bonds.

Lipophilic interactions are favorable hydrophobic-hydrophobic contacts that occur between he protein and peptide ligand. Usually, these will occur between hydrophobic amino acid side chains of the peptide buried within the pockets of the binding groove such that they are not

exposed to solvent. Exposure of the hydrophobic residues to solvent is highly unfavorable since the surrounding solvent molecules are forced to hydrogen bond with each other forming cage-like clathrate structures. The resultant decrease in entropy is highly unfavorable. Lipophilic atoms may be sulphurs which are neither polar nor hydrogen acceptors and carbon atoms which are not polar.

Van der Waal's bonds are non-specific forces found between atoms which are 3-4Å apart. They are weaker and less specific than hydrogen and electrostatic bonds. The distribution of electronic charge around an atom changes with time and, at any instant, the charge distribution is not symmetric. This transient asymmetry in electronic charge induces a similar asymmetry in neighboring atoms. The resultant attractive forces between atoms reaches a maximum at the Van der Waal's contact distance but diminishes very rapidly at about 1Å to about 2Å. Conversely, as atoms become separated by less than the contact distance, increasingly strong repulsive forces become dominant as the outer electron clouds of the atoms overlap. Although the attractive forces are relatively weak compared to electrostatic and hydrogen bonds (about 0.6 Kcal/mol), the repulsive forces in particular may be very important in determining whether a peptide ligand may bind successfully to a protein.

In one embodiment, the Böhm scoring function (SCORE1 approach) is used to estimate the binding constant. (Böhm, H.J., *J. Comput Aided Mol. Des.*, 8(3):243-256 (1994) which is hereby incorporated in its entirety). In another embodiment, the scoring function (SCORE2 approach) is used to estimate the binding affinities as an indicator of a ligand containing a T-cell epitope (Böhm, H.J., *J. Comput Aided Mol. Des.*, 12(4):309-323 (1998) which is hereby incorporated in its entirety). However, the Böhm scoring functions as described in the above references are used to estimate the binding affinity of a ligand to a protein where it is already known that the ligand successfully binds to the protein and the protein/ligand complex has had its structure solved, the solved structure being present in the Protein Data Bank ("PDB"). Therefore, the scoring function has been developed with the benefit of known positive binding data. In order to allow for discrimination between positive and negative binders, a repulsion term must be added to the equation. In addition, a more satisfactory estimate of binding energy is achieved by computing the lipophilic interactions in a pairwise manner rather than using the area based energy term of the above Böhm functions. Therefore, in a preferred embodiment, the binding energy is estimated using a modified Böhm scoring function. In the modified Böhm scoring function, the binding energy between protein and ligand (ΔG_{bind}) is estimated considering the following parameters: The reduction of binding energy due to the

overall loss of translational and rotational entropy of the ligand (ΔG_0); contributions from ideal hydrogen bonds (ΔG_{hb}) where at least one partner is neutral; contributions from unperturbed ionic interactions (ΔG_{ionic}); lipophilic interactions between lipophilic ligand atoms and lipophilic acceptor atoms (ΔG_{lipo}); the loss of binding energy due to the freezing of internal degrees of freedom in the ligand, i.e., the freedom of rotation about each C-C bond is reduced (ΔG_{rot}); the energy of the interaction between the protein and ligand (E_{VdW}). Consideration of these terms gives equation 1:

$$(\Delta G_{bind}) = (\Delta G_0) + (\Delta G_{hb} x N_{hb}) + (\Delta G_{ionic} x N_{ionic}) + (\Delta G_{lipo} x N_{lipo}) + (\Delta G_{rot} + N_{rot}) + (E_{vdw}).$$

Where N is the number of qualifying interactions for a specific term and, in one embodiment, ΔG_0, ΔG_{hb}, ΔG_{ionic}, ΔG_{lipo} and ΔG_{rot} are constants which are given the values: 5.4, -4.7, -4.7, -0.17, and 1.4, respectively.

The term N_{hb} is calculated according to equation 2:

$$N_{hb} = \sum_{h\text{-}bonds} f(\Delta R, \Delta\alpha) \times f(N_{neighb}) \times f_{pcs}$$

$f(\Delta R, \Delta\alpha)$ is a penalty function which accounts for large deviations of hydrogen bonds from ideality and is calculated according to equation 3:

$$f(\Delta R, \Delta\text{-}\alpha) = f1(\Delta R) \times f2(\Delta\alpha)$$

Where: $f1(\Delta R) = 1$ if $\Delta R <= TOL$

or $= 1 - (\Delta R - TOL)/0.4$ if $\Delta R <= 0.4 + TOL$

or $= 0$ if $\Delta R > 0.4 + TOL$

And: $f2(\Delta\alpha) = 1$ if $\Delta\alpha < 30°$

or $= 1 - (\Delta\alpha - 30)/50$ if $\Delta\alpha <= 80°$

or $= 0$ if $\Delta\alpha > 80°$

TOL is the tolerated deviation in hydrogen bond length = 0.25Å

ΔR is the deviation of the H-O/N hydrogen bond length from the ideal value = 1.9Å

$\Delta\alpha$ is the deviation of the hydrogen bond angle $\angle_{N/O\text{-}H..O/N}$ from its idealized value of 180°

$f(N_{neighb})$ distinguishes between concave and convex parts of a protein surface and therefore assigns greater weight to polar interactions found in pockets rather than those found at the protein surface. This function is calculated according to equation 4 below:

$$f(N_{neighb}) = (N_{neighb}/N_{neighb,0})^{\alpha} \text{ where } \alpha = 0.5$$

N_{neighb} is the number of non-hydrogen protein atoms that are closer than 5Å to any given protein atom.

$N_{neighb,0}$ is a constant = 25

f_{pcs} is a function which allows for the polar contact surface area per hydrogen bond and therefore distinguishes between strong and weak hydrogen bonds and its value is determined according to the following criteria:

f_{pcs}= ß when A_{polar}/N_{HB} < 10 Å^2

or f_{pcs}= 1 when A_{polar}/N_{HB} > 10 Å^2

A_{polar} is the size of the polar protein-ligand contact surface

N_{HB} is the number of hydrogen bonds

ß is a constant whose value = 1.2

For the implementation of the modified Böhm scoring function, the contributions from ionic interactions, ΔG_{ionic}, are computed in a similar fashion to those from hydrogen bonds described above since the same geometry dependency is assumed.

The term N_{lipo} is calculated according to equation 5 below:

$N_{lipo} = \sum_{lL} f(r_{lL})$

$f(r_{lL})$ is calculated for all lipophilic ligand atoms, l, and all lipophilic protein atoms, L, according to the following criteria:

$f(r_{lL}) = 1$ when $r_{lL} <= R1$ $f(r_{lL}) = (r_{lL} - R1)/(R2-R1)$ when $R2 < r_{lL} > R1$

$f(r_{lL}) = 0$ when $r_{lL} >= R2$

Where: $R1 = r_l^{vdw} + r_L^{vdw} + 0.5$

and $R2 = R1 + 3.0$

and r_l^{vdw} is the Van der Waal's radius of atom l

and r_L^{vdw} is the Van der Waal's radius of atom L

The term N_{rot} is the number of rotable bonds of the amino acid side chain and is taken to be the number of acyclic sp^3 - sp^3 and sp^3 - sp^2 bonds. Rotations of terminal $-CH_3$ or $-NH_3$ are not taken into account.

The final term, E_{VdW}, is calculated according to equation 6 below:

$E_{vdW} = \varepsilon_1\varepsilon_2((r_1^{vdw} + r_2^{vdw})^{12}/r^{12} - (r_1^{vdw} + r_2^{vdw})^6/r^6)$, where:

ε_1 and ε_2 are constants dependant upon atom identity

$r_1^{vdw} + r_2^{vdw}$ are the Van der Waal's atomic radii

r is the distance between a pair of atoms.

With regard to Equation 6, in one embodiment, the constants ε_1 and ε_2 are given the atom values: C: 0.245, N: 0.283, O: 0.316, S: 0.316, respectively (i.e. for atoms of Carbon, Nitrogen, Oxygen and Sulphur, respectively). With regards to equations 5 and 6, the Van der Waal's radii are given the atom values C: 1.85, N: 1.75, O: 1.60, S: 2.00Å.

It should be understood that all predetermined values and constants given in the equations above are determined within the constraints of current understandings of protein ligand interactions with particular regard to the type of computation being undertaken herein. Therefore, it is possible that, as this scoring function is refined further, these values and constants may change hence any suitable numerical value which gives the desired results in terms of estimating the binding energy of a protein to a ligand may be used and hence fall within the scope of the present invention.

As described above, the scoring function is applied to data extracted from the database of side-chain conformations, atom identities, and interatomic distances. For the purposes of the present description, the number of MHC Class II molecules included in this database is 42 models plus four solved structures. It should be apparent from the above descriptions that the modular nature of the construction of the computational method of the present invention means that new models can simply be added and scanned with the peptide backbone library and side-chain conformational search function to create additional data sets which can be processed by the peptide scoring function as described above. This allows for the repertoire of scanned MHC Class II molecules to easily be increased, or structures and associated data to be replaced if data are available to create more accurate models of the existing alleles.

The present prediction method can be calibrated against a data set comprising a large number of peptides whose affinity for various MHC Class II molecules has previously been experimentally determined. By comparison of calculated versus experimental data, a cut of value can be determined above which it is known that all experimentally determined T-cell epitopes are correctly predicted.

It should be understood that, although the above scoring function is relatively simple compared to some sophisticated methodologies that are available, the calculations are performed extremely rapidly. It should also be understood that the objective is not to calculate the true binding energy *per se* for each peptide docked in the binding groove of a selected MHC Class II protein. The underlying objective is to obtain comparative binding energy data as an aid to predicting the location of T-cell epitopes based on the primary structure (i.e. amino acid sequence) of a selected protein. A relatively high binding energy or a binding energy above a selected threshold value would suggest the presence of a T-cell epitope in the ligand. The ligand may then be subjected to at least one round of amino-acid substitution and the binding energy recalculated. Due to the rapid nature of the calculations, these manipulations of the peptide sequence can be performed interactively within the program's

user interface on cost-effectively available computer hardware. Major investment in computer hardware is thus not required.

It would be apparent to one skilled in the art that other available software could be used for the same purposes. In particular, more sophisticated software which is capable of docking ligands into protein binding-sites may be used in conjunction with energy minimization. Examples of docking software are: DOCK (Kuntz *et al.*, *J. Mol. Biol.*, 161:269-288 (1982)), LUDI (Böhm, H.J., *J. Comput Aided Mol. Des.*, 8:623-632 (1994)) and FLEXX (Rarey M., *et al., ISMB*, 3:300-308 (1995)). Examples of molecular modeling and manipulation software include: AMBER (Tripos) and CHARMm (Molecular Simulations Inc.). The use of these computational methods would severely limit the throughput of the method of this invention due to the lengths of processing time required to make the necessary calculations. However, it is feasible that such methods could be used as a 'secondary screen' to obtain more accurate calculations of binding energy for peptides which are found to be 'positive binders' via the method of the present invention. The limitation of processing time for sophisticated molecular mechanic or molecular dynamic calculations is one which is defined both by the design of the software which makes these calculations and the current technology limitations of computer hardware. It may be anticipated that, in the future, with the writing of more efficient code and the continuing increases in speed of computer processors, it may become feasible to make such calculations within a more manageable time-frame. Further information on energy functions applied to macromolecules and consideration of the various interactions that take place within a folded protein structure can be found in: Brooks, B.R., *et al.*, *J. Comput. Chem.*, 4:187-217 (1983) and further information concerning general protein-ligand interactions can be found in: Dauber-Osguthorpe et al., *Proteins*4(1):31-47(1988), which are incorporated herein by reference in their entirety. Useful background information can also be found, for example, in Fasman, G.D., ed., *Prediction of Protein Structure and the Principles of Protein Conformation,* Plenum Press, New York, ISBN: 0-306 4313-9.

EXAMPLE 2: *De-immunized forms of Fc-Leptin*

Leptin is a secreted signaling 146 amino acid residue protein involved in the homeostatic mechanisms maintaining adipose mass (e.g. WO 00/40615, WO 98/28427, WO 96/05309). The protein (and its antagonists) offers significant therapeutic potential for the treatment of diabetes, high blood pressure and cholesterol metabolism.

Fc-leptin is a fusion protein for which the serum half-life is profoundly improved compared to leptin itself (WO 0040615). However, certain forms of Fc-leptin, such as when the Fc is derived from human IgG1 or human IgG3, have the potential to show enhanced immunogenicity under certain circumstances, such as administration by subcutaneous injection. In a Phase I clinical trial, leptin alone was found to be at least somewhat immunogenic. The invention discloses sequences identified within the leptin primary sequence that are potential T-cell epitopes by virtue of MHC class II binding potential. This disclosure specifically pertains to the human leptin moiety containing about 146 amino acid residues. Others have provided modified leptin (US, 5,900,404; WO96/05309) but these approaches have been directed towards improvements in the commercial production of leptin, for example improved *in vitro* stability. Such teachings do not recognize the importance of T-cell epitopes to the immunogenic properties of the protein nor have been conceived to directly influence said properties in a specific and controlled way according to the scheme of the present invention. Specific Fc-leptin forms: Fcγ1-leptin, Fcγ2-leptin, both forms, preferably with linker peptide and optionally modified Fc domain having reduced affinity to Fc-receptors. Sequences to be modified in leptin are shown below:

An amino acid sequence which is part of the sequence of an immunogenically non-modified human obesity protein (leptin) and has a potential MHC class II binding activity is selected from the following group identified according to the method of the invention:

Peptide sequences in human leptin with potential human MHC class II binding activity.

VPIQKVQDDTKTL,	QKVQDDTKTLIKT,	KTLIKTIVTRIND,	TLIKTIVTRINDI,
KTIVTRINDISHT,	TIVTRINDISHTQ,	TRINDISHTQSVS,	NDISHTQSVSSKQ,
QSVSSKQKVTGLD,	SSKQKVTGLDFIP,	QKVTGLDFIPGLH,	TGLDFIPGLHPIL,
LDFIPGLHPILTL,	DFIPGLHPILTLS,	PGLHPILTLSKMD,	GLHPILTLSKMDQ,
HPILTLSKMDQTL,	PILTLSKMDQTLA,	LTLSKMDQTLAVY,	SKMDQTLAVYQQI,
QTLAVYQQILTSM,	LAVYQQILTSMPS,	AVYQQILTSMPSR,	QQILTSMPSRNVI,
QILTSMPSRNVIQ,	TSMPSRNVIQISN,	SRNVIQISNDLEN,	RNVIQISNDLENL,
NVIQISNDLENLR,	IQISNDLENLRDL,	NDLENLRDLLHVL,	LENLRDLLHVLAF,
ENLRDLLHVLAFS,	RDLLHVLAFSKSC,	DLLHVLAFSKSCH,	LHVLAFSKSCHLP,
HVLAFSKSCHLPW,	LAFSKSCHLPWAS,	CHLPWASGLETLD,	SGLETLDSLGGVL,
DSLGGVLEASGYS,	SLGGVLEASGYST,	GGVLEASGYSTEV,	SGYSTEVVALSRL,
TEVVALSRLQGSL,	EVVALSRLQGSLQ,	VALSRLQGSLQDM,	SRLQGSLQDMLWQ,
QGSLQDMLWQLDL,	GSLQDMLWQLDLS,	QDMLWQLDLSPGC	

Substitutions leading to the elimination of potential T-cell epitopes of human leptin (WT = wild type).

Residue #	WT residue	Substitutions												
3	I	A	C	D	E	G	H	K	N	P	Q	R	S	T
6	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
13	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
14	I	A	C	D	E	G	H	K	N	P	Q	R	S	T
17	I	A	C	D	E	G	H	K	N	P	Q	R	S	T
18	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
21	I	A	C	D	E	G	H	K	N	P	Q	R	S	T
24	I	A	C	D	E	G	H	K	N	P	Q	R	S	T
30	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
36	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
39	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
41	F	A	C	D	E	G	H	K	N	P	Q	R	S	T
42	I	A	C	D	E	G	H	K	N	P	Q	R	S	T
45	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
48	I	A	C	D	E	G	H	K	N	P	Q	R	S	T
49	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
51	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
54	M	A	C	D	E	G	H	K	N	P	Q	R	S	T
58	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
60	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
61	Y	A	C	D	E	G	H	K	N	P	Q	R	S	T
64	I	A	C	D	E	G	H	K	N	P	Q	R	S	T
65	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
68	M	A	C	D	E	G	H	K	N	P	Q	R	S	T
73	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
74	I	A	C	D	E	G	H	K	N	P	Q	R	S	T
76	I	A	C	D	E	G	H	K	N	P	Q	R	S	T
80	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
83	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
86	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
87	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
89	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
90	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
92	F	A	C	D	E	G	H	K	N	P	Q	R	S	T
98	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
100	W	A	C	D	E	G	H	K	N	P	Q	R	S	T
104	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
107	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
110	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
113	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
114	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
119	Y	A	C	D	E	G	H	K	N	P	Q	R	S	T
123	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
124	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
126	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
129	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
133	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
136	M	A	C	D	E	G	H	K	N	P	Q	R	S	T

Any of the above-cited peptide sequences can be used for modifying by exchanging one or more amino acids to obtain a sequence having a reduced or no immunogenicity .

EXAMPLE 3: *De-immunized forms of Fc-IL-1Ra*

The present invention provides for modified forms of an interleukin-1 receptor antagonist (IL-1Ra) with one or more T-cell epitopes removed. IL-1 is an important inflammatory and immune modulating cytokine with pleiotropic effects on a variety of tissues but may contribute to the pathology associated with rheumatoid arthritis and other diseases associated with local tissue damage. An IL-1 receptor antagonist able to inhibit the action of IL-1 has been purified and the gene cloned [Eisenburg S.P. et al (1990) *Nature*, 343: 341-346; Carter, D.B. et al (1990) *Nature*, 344: 633-637]. Others have provided IL-1Ra molecules [e.g. US 5,075,222]. Recombinant forms of this protein have therapeutic potential in disease settings where the effects of IL-1 are deleterious. However, there remains a continued need for IL-1Ra analogues with enhanced properties. Desired enhancements include alternative schemes and modalities for the expression and purification of the said therapeutic, but also and especially, improvements in the biological properties of the protein. There is a particular need for enhancement of the *in vivo* characteristics when administered to the human subject. In this regard, it is highly desired to provide IL-1Ra with reduced or absent potential to induce an immune response in the human subject. Such proteins would expect to display an increased circulation time within the human subject and would be of particular benefit in chronic or recurring disease settings such as is the case for a number of indications for IL-1Ra. The present invention provides for modified forms of IL-1Ra proteins that are expected to display enhanced properties *in vivo*. This disclosure specifically pertains a human IL-1Ra protein being of 152 amino acid residues (Eisenburg, S.P. et al (1991) *Proc. Natl. Acad. Sci. U.S.A.* 88: 5232-5236).

Specific Fc- IL-1Ra forms: Fcγ1- IL-1Ra, Fcγ2- IL-1Ra, both forms, preferably with linker peptide and optionally modified Fc domain having reduced affinity to Fc-receptors.

Peptide sequences in human interleukin-1 receptor antagonist (IL-1RA) with potential human MHC class II binding activity.

RKSSKMQAFRIWD, SKMQAFRIWDVNQ, QAFRIWDVNQKTF, FRIWDVNQKTFYL,
RIWDVNQKTFYLR, IWDVNQKTFYLRN, WDVNQKTFYLRNN, KTFYLRNNQLVAG,
TFYLRNNQLVAGY, FYLRNNQLVAGYL, LRNNQLVAGYLQG, RNNQLVAGYLQGP,
NQLVAGYLQGPNV, QLVAGYLQGPNVN, LVAGYLQGPNVNL, AGYLQGPNVNLEE,
GYLQGPNVNLEEK, PNVNLEEKIDVVP, VNLEEKIDVVPIE, EKIDVVPIEPHAL,
IDVVPIEPHALFL, DVVPIEPHALFLG, VPIEPHALFLGIH, HALFLGIHGGKMC,
ALFLGIHGGKMCL, LFLGIHGGKMCLS, LGIHGGKMCLSCV, GKMCLSCVKSGDE,
MCLSCVKSGDETR, SCVKSGDETRLQL, ETRLQLEAVNITD, TRLQLEAVNITDL,

LQLEAVNITDLSE, EAVNITDLSENRK, VNITDLSENRKQD, TDLSENRKQDKRF, ENRKQDKRFAFIR, KRFAFIRSDSGPT, FAFIRSDSGPTTS, AFIRSDSGPTTSF, TSFESAACPGWFL, SFESAACPGWFLC, PGWFLCTAMEADQ, WFLCTAMEADQPV, TAMEADQPVSLTN, QPVSLTNMPDEGV, VSLTNMPDEGVMV, TNMPDEGVMVTKF, PDEGVMVTKFYFQ, EGVMVTKFYFQED, GVMVTKFYFQEDE

Substitutions leading to the elimination of potential T-cell epitopes of human interleukin-1 receptor antagonist (IL-1RA) (WT = wild type).

Residue #	WT Residue	Substitution											
10	M	A	C	D	E	G	H	K	N	P	Q	R	S	T
13	F	A	C	D	E	G	H	K	N	P	Q	R	S	T
15	I	A	C	D	E	G	H	K	N	P	Q	R	S	T
16	W	A	C	D	E	G	H	K	N	P	Q	R	S	T
18	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
23	F	A	C	D	E	G	H	K	N	P	Q	R	S	T
24	Y	A	C	D	E	G	H	K	N	P	Q	R	S	T
25	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
30	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
31	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
34	Y	A	C	D	E	G	H	K	N	P	Q	R	S	T
35	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
40	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
42	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
46	I	A	C	D	E	G	H	K	N	P	Q	R	S	T
48	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
49	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
51	I	A	C	D	E	G	H	K	N	P	Q	R	S	T
56	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
57	F	A	C	D	E	G	H	K	N	P	Q	R	S	T
58	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
60	I	A	C	D	E	G	H	K	N	P	Q	R	S	T
65	M	A	C	D	E	G	H	K	N	P	Q	R	S	T
67	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
70	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
78	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
80	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
83	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
85	I	A	C	D	E	G	H	K	N	P	Q	R	S	T
88	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
98	F	A	C	D	E	G	H	K	N	P	Q	R	S	T
100	F	A	C	D	E	G	H	K	N	P	Q	R	S	T
101	I	A	C	D	E	G	H	K	N	P	Q	R	S	T
119	W	A	C	D	E	G	H	K	N	P	Q	R	S	T
120	F	A	C	D	E	G	H	K	N	P	Q	R	S	T
121	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
125	M	A	C	D	E	G	H	K	N	P	Q	R	S	T
131	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
133	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
136	M	A	C	D	E	G	H	K	N	P	Q	R	S	T
141	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
142	M	A	C	D	E	G	H	K	N	P	Q	R	S	T

EXAMPLE 4: *De-immunized forms of Fc-BNDF*

The present invention provides for modified forms of human brain-derived neurotrophic factor (BNDF) with one or more T-cell epitopes removed. BNDF is glycoprotein of the nerve growth factor family of proteins. The mature 119 amino acid glycoprotein is processed from a larger pre-cursor to yield a neutrophic factor that promotes the survival of neuronal cell populations [Jones K.R. & Reichardt, L.F. (1990) Proc. Natl. Acad. Sci U.S.A. 87: 8060-8064]. Others have provided modified BNDF molecules [US, 5,770,577] and approaches towards the commercial production of recombinant BNDF molecules [US, 5,986,070]. Such neuronal cells are all located either in the central nervous system or directly connected to it. Recombinant preparations of BNDF have enabled the therapeutic potential of the protein to be explored for the promotion of nerve regeneration and degenerative disease therapy.

Specific Fc- BNDF forms: Fcγ1- BNDF, Fcγ2- BNDF, both forms, preferably with linker peptide and optionally modified Fc domain having reduced affinity to Fc-receptors.

Peptide sequences in human brain-derived neutrophic factor (BDNF) with potential human MHC class II binding activity.

GELSVCDSISEWV, LSVCDSISEWVTA, DSISEWVTAADKK, SEWVTAADKKTAV,
EWVTAADKKTAVD, WVTAADKKTAVDM, KTAVDMSGGTVTV, TAVDMSGGTVTVL,
VDMSGGTVTVLEK, GTVTVLEKVPVSK, VTVLEKVPVSKGQ, TVLEKVPVSKGQL,
EKVPVSKGQLKQY, VPVSKGQLKQYFY, GQLKQYFYETKCN, KQYFYETKCNPMG,
QYFYETKCNPMGY, YFYETKCNPMGYT, NPMGYTKEGCRGI, MGYTKEGCRGIDK,
RGIDKRHWNSQCR, RHWNSQCRTTQSY, HWNSQCRTTQSYV, QSYVRALTMDSKK,
SYVRALTMDSKKR, RALTMDSKKRIGW, LTMDSKKRIGWRF, KRIGWRFIRIDTS,
IGWRFIRIDTSCV, GWRFIRIDTSCVC, WRFIRIDTSCVCT, RFIRIDTSCVCTL,
IRIDTSCVCTLTI, IDTSCVCTLTIKR

Substitutions leading to the elimination of potential T-cell epitopes of human brain-derived neutrophic factor (BDNF) (WT = wild type).

Residue #	WT Residue	Substitution												
10	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
16	I	A	C	D	E	G	H	K	N	P	Q	R	S	T
20	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
29	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
31	M	A	C	D	E	G	H	K	N	P	Q	R	S	T
36	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
38	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
39	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
42	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
44	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
49	L	A	C	D	E	G	H	K	N	P	Q	R	S	T

Residue #	WT Residue	Substitution												
52	Y	A	C	D	E	G	H	K	N	P	Q	R	S	T
53	F	A	C	D	E	G	H	K	N	P	Q	R	S	T
54	Y	A	C	D	E	G	H	K	N	P	Q	R	S	T
61	M	A	C	D	E	G	H	K	N	P	Q	R	S	T
63	Y	A	C	D	E	G	H	K	N	P	Q	R	S	T
71	I	A	C	D	E	G	H	K	N	P	Q	R	S	T
76	W	A	C	D	E	G	H	K	N	P	Q	R	S	T
86	Y	A	C	D	E	G	H	K	N	P	Q	R	S	T
87	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
90	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
92	M	A	C	D	E	G	H	K	N	P	Q	R	S	T
98	I	A	C	D	E	G	H	K	N	P	Q	R	S	T
100	W	A	C	D	E	G	H	K	N	P	Q	R	S	T
102	F	A	C	D	E	G	H	K	N	P	Q	R	S	T
103	I	A	C	D	E	G	H	K	N	P	Q	R	S	T
105	I	A	C	D	E	G	H	K	N	P	Q	R	S	T

EXAMPLE 5: *De-immunized forms of Fc-EPO*

The present invention provides for modified forms of human erythropioetin (EPO) with one or more T-cell epitopes removed. EPO is a 165 amino acid residues glycoprotein hormone involved in the maturation of erythroid progenitor cells into erythrocytes. Naturally occurring EPO is produced by the liver during foetal life and by the kidney of adults and circulates in the blood to stimulate production of red blood cells in bone marrow. Anaemia is almost invariably a consequence of renal failure due to decreased production of EPO from the kidney. Recombinant EPO is used as an effective treatment of anaemia resulting from chronic renal failure.

Recombinant EPO (expressed in mammalian cells) having the amino acid sequence 1-165 of human erythropoietin [Jacobs, K. et al (1985) Nature, 313: 806-810; Lin, F.-K. et al (1985) Proc. Natl. Acad. Sci. U.S.A. 82:7580-7585] contains three N-linked and one O-linked oligosaccharide chains each containing terminal sialic acid residues. The latter are significant in enabling EPO to evade rapid clearance from the circulation by the hepatic asialoglycoprotein binding protein.

Non-de-immunized Fc-EPO is known e.g. from WO 99/58662, WO 99/02709.

Specific Fc- EPO forms: Fcγ1- EPO, Fcγ2- EPO, both forms, preferably with linker peptide and optionally modified Fc domain having reduced affinity to Fc-receptors. The EPO may be glycosylated, partially glycosylated or has a modified glycosylation pattern.

Peptide sequences in human erythropoietin (EPO) with potential human MHC class II binding activity.

PRLICDSRVLERY, RLICDSRVLERYL, ICDSRVLERYLLE, CDSRVLERYLLEA, SRVLERYLLEAKE, RVLERYLLEAKEA, LERYLLEAKEAEN, ERYLLEAKEAENI, RYLLEAKEAENIT, YLLEAKEAENITT, LEAKEAENITTGC, KEAENITTGCAEH, ENITTGCAEHCSL, CSLNENITVPDTK, NENITVPDTKVNF, ENITVPDTKVNFY, NITVPDTKVNFYA, ITVPDTKVNFYAW, TKVNFYAWKRMEV, VNFYAWKRMEVGQ, NFYAWKRMEVGQQ, YAWKRMEVGQQAV, KRMEVGQQAVEVW, RMEVGQQAVEVWQ, MEVGQQAVEVWQG, QAVEVWQGLALLS, AVEVWQGLALLSE, VEVWQGLALLSEA, EVWQGLALLSEAV, VWQGLALLSEAVL, WQGLALLSEAVLR, QGLALLSEAVLRG, LALLSEAVLRGQA, ALLSEAVLRGQAL, LSEAVLRGQALLV, SEAVLRGQALLVN, EAVLRGQALLVNS, AVLRGQALLVNSS, QALLVNSSQPWEP, ALLVNSSQPWEPL, LLVNSSQPWEPLQ, QPWEPLQLHVDKA, EPLQLHVDKAVSG, LQLHVDKAVSGLR, LHVDKAVSGLRSL, KAVSGLRSLTTLL, SGLRSLTTLLRAL, RSLTTLLRALGAQ, SLTTLLRALGAQK, TTLLRALGAQKEA, TLLRALGAQKEAI, RALGAQKEAISPP, AQKEAISPPDAAS, EAISPPDAASAAP, SPPDAASAAPLRT, ASAAPLRTITADT, APLRTITADTFRK, RTITADTFRKLFR, TITADTFRKLFRV, DTFRKLFRVYSNF, RKLFRVYSNFLRG, KLFRVYSNFLRGK, FRVYSNFLRGKLK, RVYSNFLRGKLKL, YSNFLRGKLKLYT, SNFLRGKLKLYTG, NFLRGKLKLYTGE, RGKLKLYTGEACR, GKLKLYTGEACRT, LKLYTGEACRTGD, KLYTGEACRTGDR

Substitutions leading to the elimination of potential T-cell epitopes of human erythropoietin (EPO) (WT = wild type).

Residue #	WT residue	Substitutions											
5	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
6	I	A	C	D	E	G	H	K	N	P	Q	R	S	T
11	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
12	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
15	Y	A	C	D	E	G	H	K	N	P	Q	R	S	T
16	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
17	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
25	I	A	C	D	E	G	H	K	N	P	Q	R	S	T
35	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
39	I	A	C	D	E	G	H	K	N	P	Q	R	S	T
41	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
46	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
48	F	A	C	D	E	G	H	K	N	P	Q	R	S	T
49	Y	A	C	D	E	G	H	K	N	P	Q	R	S	T
51	W	A	C	D	E	G	H	K	N	P	Q	R	S	T
54	M	A	C	D	E	G	H	K	N	P	Q	R	S	T
56	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
61	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
63	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
64	W	A	C	D	E	G	H	K	N	P	Q	R	S	T
67	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
69	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
70	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
74	V	A	C	D	E	G	H	K	N	P	Q	R	S	T

Residue #	WT residue	Substitutions											
75	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
80	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
81	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
82	N	A	C	D	E	G	H	K	N	P	Q	R	S	T
88	W	A	C	D	E	G	H	K	N	P	Q	R	S	T
91	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
93	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
95	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
99	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
102	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
105	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
108	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
109	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
112	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
119	I	A	C	D	E	G	H	K	N	P	Q	R	S	T
130	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
133	I	A	C	D	E	G	H	K	N	P	Q	R	S	T
138	F	A	C	D	E	G	H	K	N	P	Q	R	S	T
141	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
142	F	A	C	D	E	G	H	K	N	P	Q	R	S	T
144	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
145	Y	A	C	D	E	G	H	K	N	P	Q	R	S	T
148	F	A	C	D	E	G	H	K	N	P	Q	R	S	T
149	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
153	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
155	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
156	Y	A	C	D	E	G	H	K	N	P	Q	R	S	T

EXAMPLE 6: *De-immunized forms of G-CSF*

The present invention provides for modified forms of human granulocyte colony stimulating factor (G-CSF) with one or more T-cell epitopes removed. G-CSF is an important haemopoietic cytokine currently used in treatment of indications where an increase in blood neutrophils will provide benefits. These include cancer therapy, various infectious diseases and related conditions such as sepsis. G-CSF is also used alone, or in combination with other compounds and cytokines in the *ex vivo* expansion of haemopoeitic cells for bone marrow transplantation.

Two forms of human G-CSF are commonly recognized for this cytokine. One is a protein of 177 amino acids, the other a protein of 174 amino acids [Nagata et al. (1986), EMBO J. 5: 575-581], the 174 amino acid form has been found to have the greatest specific *in vivo* biological activity. Recombinant DNA techniques have enabled the production of commercial scale quantities of G-CSF exploiting both eukaryotic and prokaryotic host cell expression systems. This disclosure specifically pertains to both recognized forms of the human G-CSF protein being the 177 amino acid species and the 174 amino acid species.

Other polypeptide analogues and peptide fragments of G-CSF have been previously disclosed, including forms modified by site-specific amino acid substitutions and or by modification by chemical adducts. Thus US 4,810,643 discloses analogues with the particular Cys residues replaced with another amino acid, and G-CSF with an Ala residue in the first (N-terminal) position. EP 0 335 423 discloses the modification of at least one amino group in a polypeptide having G-CSF activity. EP 0 272 703 discloses G-CSF derivatives having amino acid substituted or deleted near the N-terminus. EP 0 459 630 discloses G-CSF derivatives in which Cys 17 and Asp 27 are replaced by Ser residues. EP 0 243 153 discloses G-CSF modified by inactivating at least one yeast KEX2 protease processing site for increased yield in recombinant production and US 4,904,584 discloses lysine altered proteins. WO 90/12874 discloses further Cys altered variants and Australian patent document AU 10948/92 discloses the addition of amino acids to either terminus of a G-CSF molecule for the purpose of aiding in the folding of the molecule after prokaryotic expression. A further Australian document; AU 76380/91, discloses G-CSF variants at positions 50-56 of the G-CSF 174 amino acid form, and positions 53-59 of the 177 amino acid form. Additional changes at particular His residues were also disclosed.

Non-deimmunized Fc-G-CSF is known e.g. from WO 99/58662. Specific Fc- G-CSF forms: Fcγ1- G-CSF, Fcγ2- G-CSF, both forms, preferably with linker peptide and optionally modified Fc domain having reduced affinity to Fc-receptors.

Peptide sequences in human granulocyte colony stimulating factor (G-CSF) with potential human MHC class II binding activity.

TPLGPASSLPQSF,	SSLPQSFLLKCLE,	QSFLLKCLEQVRK,	SFLLKCLEQVRKI,
FLLKCLEQVRKIQ,	KCLEQVRKIQGDG,	EQVRKIQGDGAAL,	RKIQGDGAALQEK,
AALQEKLVSECAT,	EKLVSECATYKLC,	KLVSECATYKLCH,	AALQEKLCATYKL,
EKLCATYKLCHPE,	ATYKLCHPEELVL,	YKLCHPEELVLLG,	EELVLLGHSLGIP,
ELVLLGHSLGIPW,	HSLGIPWAPLSSC,	IPWAPLSSCPSQA,	APLSSCPSQALQL,
QALQLAGCLSQLH,	GCLSQLHSGLFLY,	SQLHSGLFLYQGL,	SGLFLYQGLLQAL,
GLFLYQGLLQALE,	LFLYQGLLQALEG,	FLYQGLLQALEGI,	QGLLQALEGISPE,
GLLQALEGISPEL,	QALEGISPELGPT,	EGISPELGPTLDT,	PTLDTLQLDVADF,
DTLQLDVADFATT,	LQLDVADFATTIW,	LDVADFATTIWQQ,	TTIWQQMEELGMA,
TIWQQMEELGMAP,	QQMEELGMAPALQ,	EELGMAPALQPTQ,	LGMAPALQPTQGA,
PALQPTQGAMPAF,	GAMPAFASAFQRR,	PAFASAFQRRAGG,	SAFQRRAGGVLVA,
GGVLVASHLQSFL,	GVLVASHLQSFLE,	VLVASHLQSFLEV,	SHLQSFLEVSYRV,
QSFLEVSYRVLRH,	SFLEVSYRVLRHL,	LEVSYRVLRHLAQ	

Substitutions leading to the elimination of potential T-cell epitopes of human granulocyte colony stimulating factor (G-CSF) (WT = wild type).

Residue #	WT Residue	Substitution												
3	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
9	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
14	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
15	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
18	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
21	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
24	I	A	C	D	E	G	H	K	N	P	Q	R	S	T
31	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
35	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
39	Y	A	C	D	E	G	H	K	N	P	Q	R	S	T
41	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
47	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
48	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
49	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
50	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
54	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
56	I	A	C	D	E	G	H	K	N	P	Q	R	S	T
58	W	A	C	D	E	G	H	K	N	P	Q	R	S	T
61	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
69	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
71	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
75	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
78	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
82	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
83	F	A	C	D	E	G	H	K	N	P	Q	R	S	T
84	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
85	Y	A	C	D	E	G	H	K	N	P	Q	R	S	T
88	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
89	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
92	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
95	I	A	C	D	E	G	H	K	N	P	Q	R	S	T
99	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
103	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
106	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
108	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
110	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
113	F	A	C	D	E	G	H	K	N	P	Q	R	S	T
117	I	A	C	D	E	G	H	K	N	P	Q	R	S	T
118	W	A	C	D	E	G	H	K	N	P	Q	R	S	T
121	M	A	C	D	E	G	H	K	N	P	Q	R	S	T
124	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
130	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
137	M	A	C	D	E	G	H	K	N	P	Q	R	S	T
140	F	A	C	D	E	G	H	K	N	P	Q	R	S	T
144	F	A	C	D	E	G	H	K	N	P	Q	R	S	T
151	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
152	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
153	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
157	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
160	F	A	C	D	E	G	H	K	N	P	Q	R	S	T
161	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
163	V	A	C	D	E	G	H	K	N	P	Q	R	S	T

EXAMPLE 7: *De-immunized forms of KGF*

The present invention provides for modified forms of human keratinocyte growth factor (KGF) with one or more T-cell epitopes removed. KGF is a member of the fibroblast growth factor (FGF) / heparin-binding growth factor family of proteins. It is a secreted glycoprotein expressed predominantly in the lung, promoting wound healing by stimulating the growth of keratinocytes and other epithelial cells [Finch et al (1989), Science 24: 752-755; Rubin et al (1989), Proc. Natl. Acad. Sci. U.S.A. 86: 802-806]. The mature (processed) form of the glycoprotein comprises 163 amino acid residues and may be isolated from conditioned media following culture of particular cell lines [Rubin et al, (1989) *ibid.*], or produced using recombinant techniques [Ron et al (1993) J. Biol. Chem. 268: 2984-2988]. The protein is of therapeutic value for the stimulation of epithelial cell growth in a number of significant disease and injury repair settings. This disclosure specifically pertains the human KGF protein being the mature (processed) form of 163 amino acid residues. Others have also provided KGF molecules [e.g. US, 6,008,328; WO90/08771;] including modified KGF [Ron et al (1993) *ibid*; WO9501434] but

Specific Fc- KGF forms: Fcγ1- KGF, Fcγ2- KGF, both forms, preferably with linker peptide and optionally modified Fc domain having reduced affinity to Fc-receptors.

Peptide sequences in human keratinocyte growth factor (KGF) with potential human MHC class II binding activity.

NDMTPEQMATNVN, DMTPEQMATNVNC, EQMATNVNCSSPE, TNVNCSSPERHTR,
RSYDYMEGGDIRV, YDYMEGGDIRVRR, DYMEGGDIRVRRL, GDIRVRRLFCRTQ,
IRVRRLFCRTQWY, RRLFCRTQWYLRI, RLFCRTQWYLRID, TQWYLRIDKRGKV,
QWYLRIDKRGKVK, WYLRIDKRGKVKG, LRIDKRGKVKGTQ, GKVKGTQEMKNNY,
QEMKNNYNIMEIR, NNYNIMEIRTVAV, YNIMEIRTVAVGI, NIMEIRTVAVGIV,
MEIRTVAVGIVAI, RTVAVGIVAIKGV, VAVGIVAIKGVES, VGIVAIKGVESEF,
VAIKGVESEFYLA, KGVESEFYLAMNK, SEFYLAMNKEGKL, EFYLAMNKEGKLY,
FYLAMNKEGKLYA, LAMNKEGKLYAKK, GKLYAKKECNEDC, KLYAKKECNEDCN,
CNFKELILENHYN, KELILENHYNTYA, ELILENHYNTYAS, LILENHYNTYASA,
NHYNTYASAKWTH, NTYASAKWTHNGG, AKWTHNGGEMFVA, GEMFVALNQKGIP,
EMFVALNQKGIPV, FVALNQKGIPVRG, VALNQKGIPVRGK, KGIPVRGKKTKKE,
IPVRGKKTKKEQK, KTKKEQKTAHFLP

Substitutions leading to the elimination of potential T-cell epitopes of human keratinocyte growth factor (KGF) (WT = wild type).

Residue #	WT residue	Substitution												
5	M	A	C	D	E	G	H	K	N	P	Q	R	S	T
10	M	A	C	D	E	G	H	K	N	P	Q	R	S	T
14	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
26	Y	A	C	D	E	G	H	K	N	P	Q	R	S	T
28	Y	A	C	D	E	G	H	K	N	P	Q	R	S	T
29	M	A	C	D	E	G	H	K	N	P	Q	R	S	T
34	I	A	C	D	E	G	H	K	N	P	Q	R	S	T
36	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
39	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
40	F	A	C	D	E	G	H	K	N	P	Q	R	S	T
45	W	A	C	D	E	G	H	K	N	P	Q	R	S	T
46	Y	A	C	D	E	G	H	K	N	P	Q	R	S	T
47	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
49	I	A	C	D	E	G	H	K	N	P	Q	R	S	T
55	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
61	M	A	C	D	E	G	H	K	N	P	Q	R	S	T
65	Y	A	C	D	E	G	H	K	N	P	Q	R	S	T
67	I	A	C	D	E	G	H	K	N	P	Q	R	S	T
68	M	A	C	D	E	G	H	K	N	P	Q	R	S	T
70	I	A	C	D	E	G	H	K	N	P	Q	R	S	T
73	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
75	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
77	I	A	C	D	E	G	H	K	N	P	Q	R	S	T
78	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
80	I	A	C	D	E	G	H	K	N	P	Q	R	S	T
83	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
87	F	A	C	D	E	G	H	K	N	P	Q	R	S	T
88	Y	A	C	D	E	G	H	K	N	P	Q	R	S	T
89	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
91	M	A	C	D	E	G	H	K	N	P	Q	R	S	T
97	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
98	Y	A	C	D	E	G	H	K	N	P	Q	R	S	T
109	F	A	C	D	E	G	H	K	N	P	Q	R	S	T
112	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
113	I	A	C	D	E	G	H	K	N	P	Q	R	S	T
114	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
118	Y	A	C	D	E	G	H	K	N	P	Q	R	S	T
121	Y	A	C	D	E	G	H	K	N	P	Q	R	S	T
126	W	A	C	D	E	G	H	K	N	P	Q	R	S	T
133	M	A	C	D	E	G	H	K	N	P	Q	R	S	T
134	F	A	C	D	E	G	H	K	N	P	Q	R	S	T
135	V	A	C	D	E	G	H	K	N	P	Q	R	S	T
137	L	A	C	D	E	G	H	K	N	P	Q	R	S	T
142	I	A	C	D	E	G	H	K	N	P	Q	R	S	T
144	V	A	C	D	E	G	H	K	N	P	Q	R	S	T

EXAMPLE 8: *De-immunized sTNF-R(I) and sTNF Inhibitor within corresponding Fc fusions*

Fc-sTNF-R(I) and Fc-sTNF Inhibitor are fusion proteins in which the serum half-life is extended compared to sTNF-R(I) and sTNF Inhibitor itself. However, certain forms of Fc-TNF-RI, such as when the Fc is derived from human IgG1 or human IgG3, have the potential

to show enhanced immunogenicity under certain circumstances, such as administration by subcutaneous injection. The present invention provides for modified forms of a soluble tumor necrosis factor receptor type I (sTNF-RI) with one or more T-cell epitopes removed.

The sTNF-RI (soluble tumor necrosis factor receptor type I) is a derivative of the human tumor necrosis factor receptor described previously [Gray, P.W. et al (1990) *Proc. Nat. Acad. Sci. U.S.A.* 87: 7380-7384; Loetschere, H. et al, (1990) *Cell* 61: 351-359; Schall, T.J. et al (1990) *Cell* 61: 361-370], comprising the extracellular domain of the intact receptor and exhibiting an approximate molecular weight of 30KDa. Additional soluble TNF inhibitors and in particular a 40KDa form are also known [US 6,143,866]. The soluble forms are able to bind tumor necrosis factor alpha with high affinity and inhibit the cytotoxic activity of the cytokine *in vitro*. Recombinant preparations of sTNF-RI are of significant therapeutic value for the treatment of diseases where an excess level of tumor necrosis factor is causing a pathogenic effect. Indications such as cachexia, sepsis and autoimmune disorders including, and in particular, rheumatoid arthritis and others may be targeted by such therapeutic preparations of sTNF-RI. Others including Brewer et al., US, 6,143,866, have provided modified sTNF-RI molecules

Peptide sequences in a human 30KDa sTNF-RI with potential human MHC class II binding activity:

DSVCPQGKYIHPQ, KYIHPQNNSICCT, NSICCTKCHKGTY, TYLYNDCPGPGQD,
YLYNDCPGPGQDT, NHLRHCLSCSKCR, HCLSCSKCRKEMG, KEMGQVEISSCTV,
GQVEISSCTVDRD, VEISSCTVDRDTV, CTVDRDTVCGCRK, DTVCGCRKNQYRH,
NQYRHYWSENLFQ, RHYWSENLFQCFN, HYWSENLFQCFNC, ENLFQCFNCSLCL,
NLFQCFNCSLCLN, QCFNCSLCLNGTV, CSLCLNGTVHLSC, LCLNGTVHLSCQE,
GTVHLSCQEKQNT, VHLSCQEKQNTVC, EKQNTVCTCHAGF, NTVCTCHAGFFLR,
GFFLRENECVSCS, FFLRENECVSCSN, ECVSCSNCKKSLE, KSLECTKLCLPQI,
TKLCLPQIENVKG, LCLPQIENVKGTE, PQIENVKGTEDSG, SGTTVLLPLVIFF

Peptide sequences in a human 40KDa sTNF inhibitor with potential human MHC class II binding activity.

TPYAPEPGSTCRL, CRLREYYDQTAQM, REYYDQTAQMCCS, EYYDQTAQMCCSK,
AQMCCSKCSPGQH, KCSPGQHAKVFCT, AKVFCTKTSDTVC, KVFCTKTSDTVCD,
STYTQLWNWVPEC, TQLWNWVPECLSC, QLWNWVPECLSCG, NWVPECLSCGSRC,
ECLSCGSRCSSDQ, SRCSSDQEVTQAC, QEVTQACTREQNR, QNRICTCRPGWYC,
NRICTCRPGWYCA, PGWYCALSKQEGC, GWYCALSKQEGCR, CALSKQEGCRLCA,
APLRKCRPGFGVA, PGFGVARPGTETS, FGVARPGTETSDV, SDVVCKPCAPGTF,
GTFSNTTSSTDIC, TDICRPHQICNVV, HQICNVVAIPGNA, ICNVVAIPGNASR,
CNVVAIPGNASRD, NVVAIPGNASRDA, VAIPGNASRDAVC, DAVCTSTTTPTRS,

TRSMAPGAVHLPQ, RSMAPGAVHLPQP, VHLPQPVSTRSQH, QPVSTRSQHTQPT,
PEPSTAPSTSFLL, SFLLPMGPSPPAE, FLLPMGPSPPAEG

EXAMPLE 9 (soluble TNF-R2):

Fc-sTNF-R2 is a fusion proteins in which the serum half-life is extended compared to sTNF-R2 itself. However, certain forms of Fc-TNF-R2, such as when the Fc is derived from human IgG1 or human IgG3, have the potential to show enhanced immunogenicity under certain circumstances, such as administration by subcutaneous injection.

Soluble tumor necrosis factor receptor 2 (sTNF-R2) is a derivative of the human tumor necrosis factor receptor 2 described previously [Smith, C.A. et al (1990) *Science* 248: 1019-1023; Kohno, T. et al (1990) *Proc. Nat. Acad. Sci. U.S.A.* 87: 8331-8335; Beltinger, C.P. et al (1996) *Genomics* 35:94-100] comprising the extracellular domain of the intact receptor. The soluble forms are able to bind tumour necrosis factor with high affinity and inhibit the cytotoxic activity of the cytokine *in vitro*. Recombinant preparations of sTNF-R2 are of significant therapeutic value for the treatment of diseases where an excess level of tumour necrosis factor is causing a pathogenic effect. A particular recombinant preparation termed ethanercept has gained clinical approval for the treatment of rheumatoid arthritis and this and other similar agents may be of value in the treatment of other indications such as cachexia, sepsis and autoimmune disorders. Ethanercept is a dimeric fusion protein comprising the extracellular domain of the human TNFR2 molecule in combination with the Fc domain of the human IgG1 molecule. The dimeric molecule comprises 934 amino acids [US,5,395,760; US,5,605,690; US,5,945,397].

Peptide sequences in the TNF binding domain of the human TNFR2 protein with potential human MHC class II binding activity are:

TPYAPEPGSTCRL, CRLREYYDQTAQM, REYYDQTAQMCCS, EYYDQTAQMCCSK,
AQMCCSKCSPGQH, KCSPGQHAKVFCT, AKVFCTKTSDTVC, KVFCTKTSDTVCD,
STYTQLWNWVPEC, TQLWNWVPECLSC, QLWNWVPECLSCG, NWVPECLSCGSRC,
ECLSCGSRCSSDQ, SRCSSDQEVTQAC, QEVTQACTREQNR, QNRICTCRPGWYC,
NRICTCRPGWYCA, PGWYCALSKQEGC, GWYCALSKQEGCR, CALSKQEGCRLCA,
APLRKCRPGFGVA, PGFGVARPGTETS, FGVARPGTETSDV, SDVVCKPCAPGTF,
GTFSNTTSSTDIC, TDICRPHQICNVV, HQICNVVAIPGNA, ICNVVAIPGNASR,
CNVVAIPGNASRD, NVVAIPGNASRDA, VAIPGNASRDAVC, DAVCTSTTTPTRS,
TRSMAPGAVHLPQ, RSMAPGAVHLPQP, VHLPQPVSTRSQH, QPVSTRSQHTQPT,
PEPSTAPSTSFLL, SFLLPMGPSPPAE, FLLPMGPSPPAEG

EXAMPLE 10: *non-natural forms of Beta-Glucocerebrosidase (ß-GCR)*

Fc – ß-GCR is a fusion proteins in which the serum half-life is extended compared to the ß-GCR itself. However, certain forms of Fc – ß-GCR, such as when the Fc is derived from human IgG1 or human IgG3, have the potential to show enhanced immunogenicity under certain circumstances, such as administration by subcutaneous injection. The present invention provides for modified forms of human GCR, preferably Fc-ß-GCR, with one or more T-cell epitopes removed.

Beta-Glucocerebrosidase (b-D-glucosyl-N-acylsphingosine glucohydrolase, E.C. 3.2.1.45) is a monomeric glycoprotein of 497 amino acid residues. The enzyme catalyses the hydrolysis of the glycolipid glucocerebroside to glucose and ceramide. Deficiency in GCR activity results in a lysosomal storage disease referred to as Gaucher disease. The disease is characterised by the accumulation of glucocerebroside engorged tissue macrophages that accumulate in the liver, spleen, bone marrow and other organs. The disease has varying degrees of severity from type 1 disease with haematologic problems but no neuronal involvement, to type 2 disease manifesting early after birth with extensive neuronal involvement and is universally progressive and fatal within 2 years of age. Type 3 disease is also recognised in some classifications and also shows neurologic involvement. Previously the only useful therapy for Gaucher disease has been administration of GCR derived from human placenta (known as alglucerase) but more recently pharmaceutical preparations of recombinant GCR ("Ceredase" and "Cerezyme") have shown efficacy in the treatment of type I disease [Niederau, C. et al (1998) *Eur. J. Med. Res.* 3: 25-30].

According to the invention, the particular commercial forms of glucocerebrosidase are examined at predicted to be particularly immunogenic because these forms are engineered to have a high mannose oligosaccharide. Upon administration of such a protein, such as Ceredase or Cerezyme, the non-natural protein is preferentially bound by mannose receptors on antigen-presenting cells such as macrophages or dendritic cells. The non-natural protein is then taken up, a portion is degraded into peptides, and the peptides presented through MHC Class II to the T-cell receptor. By mutating the glucocerebrosidase sequence such that derived peptides cannot bind to MHC Class II, immunogenicity is reduced.

Others have provided GCR molecules including modified GCR [US,5,236,838] but this teaching does not recognize the importance of T-cell epitopes to the immunogenic properties of the protein.

Peptide sequences in human ß-GCR with potential human MHC class II binding activity are:

PCIPKSFGYSSVV, KSFGYSSVVCVCN, FGYSSVVCVCNAT, SSVVCVCNATYCD, SVVCVCNATYCDS,
VCVCNATYCDSFD, ATYCDSFDPPTFP, DSFDPPTFPALGT, PTFPALGTFSRYE, PALGTFSRYESTR,
GTFSRYESTRSGR, SRYESTRSGRRME, GRRMELSMGPIQA, RRMELSMGPIQAN, RMELSMGPIQANH,
MELSMGPIQANHT, LSMGPIQANHTGT, MGPIQANHTGTGL, GPIQANHTGTGLL, TGLLLTLQPEQKF,
GLLLTLQPEQKFQ, LLLTLQPEQKFQK, LTLQPEQKFQKVK, TLQPEQKFQKVKG, PEQKFQKVKGFGG,
QKFQKVKGFGGAM, QKVKGFGGAMTDA, KGFGGAMTDAAAL, GFGGAMTDAAALN, GAMTDAAALNILA,
AMTDAAALNILAL, MTDAAALNILALS, AALNILALSPPAQ, ALNILALSPPAQN, LNILALSPPAQNL,
NILALSPPAQNLL, LALSPPAQNLLLK, ALSPPAQNLLLKS, PAQNLLLKSYFSE, AQNLLLKSYFSEE,
QNLLLKSYFSEEG, NLLLKSYFSEEGI, LLLKSYFSEEGIG, KSYFSEEGIGYNI, SYFSEEGIGYNII,
FSEEGIGYNIIRV, EGIGYNIIRVPMA, GIGYNIIRVPMAS, IGYNIIRVPMASC, YNIIRVPMASCDF,
NIIRVPMASCDFS, IIRVPMASCDFSI, IRVPMASCDFSIR, VPMASCDFSIRTY, PMASCDFSIRTYT,
SCDFSIRTYTYAD, CDFSIRTYTYADT, FSIRTYTYADTPD, RTYTYADTPDDFQ, TYTYADTPDDFQL,
YTYADTPDDFQLH, ADTPDDFQLHNFS, PDDFQLHNFSLPE, DDFQLHNFSLPEE, FQLHNFSLPEEDT,
HNFSLPEEDTKLK, FSLPEEDTKLKIP, SLPEEDTKLKIPL, EEDTKLKIPLIHR, TKLKIPLIHRALQ,
KLKIPLIHRALQL, LKIPLIHRALQLA, IPLIHRALQLAQR, PLIHRALQLAQRP, HRALQLAQRPVSL,
RALQLAQRPVSLL, ALQLAQRPVSLLA, LQLAQRPVSLLAS, RPVSLLASPWTSP, PVSLLASPWTSPT,
VSLLASPWTSPTW, SLLASPWTSPTWL, SPWTSPTWLKTNG, TSPTWLKTNGAVN, PTWLKTNGAVNGK,
TWLKTNGAVNGKG, GAVNGKGSLKGQP, GSLKGQPGDIYHQ, GDIYHQTWARYFV, DIYHQTWARYFVK,
QTWARYFVKFLDA, WARYFVKFLDAYA, ARYFVKFLDAYAE, RYFVKFLDAYAEH, YFVKFLDAYAEHK,
FVKFLDAYAEHKL, VKFLDAYAEHKLQ, KFLDAYAEHKLQF, DAYAEHKLQFWAV, YAEHKLQFWAVTA,
HKLQFWAVTAENE, LQFWAVTAENEPS, QFWAVTAENEPSA, FWAVTAENEPSAG, WAVTAENEPSAGL,
VTAENEPSAGLLS, PSAGLLSGYPFQC, AGLLSGYPFQCLG, GLLSGYPFQCLGF, SGYPFQCLGFTPE,
YPFQCLGFTPEHQ, QCLGFTPEHQRDF, LGFTPEHQRDFIA, FTPEHQRDFIARD, RDFIARDLGPTLA,
DFIARDLGPTLAN, RDLGPTLANSTHH, LGPTLANSTHHNV, PTLANSTHHNVRL, HNVRLLMLDDQRL,
VRLLMLDDQRLLL, RLLMLDDQRLLLP, LLMLDDQRLLLPH, LMLDDQRLLLPHW, DDQRLLLPHWAKV,
DQRLLLPHWAKVV, QRLLLPHWAKVVL, RLLLPHWAKVVLT, LLLPHWAKVVLTD, PHWAKVVLTDPEA,
WAKVVLTDPEAAK, AKVVLTDPEAAKY, KVVLTDPEAAKYV, VVLTDPEAAKYVH, EAAKYVHGIAVHW,
AKYVHGIAVHWYL, KYVHGIAVHWYLD, YVHGIAVHWYLDF, HGIAVHWYLDFLA, IAVHWYLDFLAPA,
VHWYLDFLAPAKA, HWYLDFLAPAKAT, WYLDFLAPAKATL, LDFLAPAKATLGE, DFLAPAKATLGET,
AKATLGETHRLFP, ATLGETHRLFPNT, GETHRLFPNTMLF, ETHRLFPNTMLFA, THRLFPNTMLFAS,
HRLFPNTMLFASE, RLFPNTMLFASEA, FPNTMLFASEACV, NTMLFASEACVGS, TMLFASEACVGSK,
MLFASEACVGSKF, ACVGSKFWEQSVR, GSKFWEQSVRLGS, SKFWEQSVRLGSW, KFWEQSVRLGSWD,
QSVRLGSWDRGMQ, VRLGSWDRGMQYS, RLGSWDRGMQYSH, GSWDRGMQYSHSI, WDRGMQYSHSIIT,
RGMQYSHSIITNL, MQYSHSIITNLLY, QYSHSIITNLLYH, YSHSIITNLLYHV, HSIITNLLYHVVG,
SIITNLLYHVVGW, TNLLYHVVGWTDW, NLLYHVVGWTDWN, LLYHVVGWTDWNL, YHVVGWTDWNLAL,
HVVGWTDWNLALN, VVGWTDWNLALNP, VGWTDWNLALNPE, TDWNLALNPEGGP, WNLALNPEGGPNW,
LALNPEGGPNWVR, PNWVRNFVDSPII, NWVRNFVDSPIIV, RNFVDSPIIVDIT, NFVDSPIIVDITK,
SPIIVDITKDTFY, PIIVDITKDTFYK, IIVDITKDTFYKQ, VDITKDTFYKQPM, DTFYKQPMFYHLG,
TFYKQPMFYHLGH, QPMFYHLGHFSKF, PMFYHLGHFSKFI, MFYHLGHFSKFIP, YHLGHFSKFIPEG,

GHFSKFIPEGSQR, SKFIPEGSQRVGL, KFIPEGSQRVGLV, IPEGSQRVGLVAS, QRVGLVASQKNDL, VGLVASQKNDLDA, GLVASQKNDLDAV, SQKNDLDAVALMH, NDLDAVALMHPDG, DAVALMHPDGSAV, VALMHPDGSAVVV, ALMHPDGSAVVVV, SAVVVVLNRSSKD, AVVVVLNRSSKDV, VVVVLNRSSKDVP, VVVLNRSSKDVPL, VVLNRSSKDVPLT, KDVPLTIKDPAVG, VPLTIKDPAVGFL, PLTIKDPAVGFLE, LTIKDPAVGFLET, PAVGFLETISPGY, VGFLETISPGYSI, GFLETISPGYSIH, FLETISPGYSIHT, ETISPGYSIHTYL, PGYSIHTYLWHRQ, PGYSIHTYLWRRQ

EXAMPLE 11: *De-immunized forms of Fc-IL2*

Non-deimmunized Fc-IL2 was described e.g. in WO 96/08570. Specific de-immunized Fc-IL2 forms: Fcγ1- IL2 , Fcγ2- IL2 , both forms, preferably with linker peptide and optionally modified Fc domain having reduced affinity to Fc-receptors.

EXAMPLE 12: *De-immunized forms Fc-IL12*

Non-deimmunized Fc- IL12 was described e.g. in WO 99/29732. Specific de-immunized Fc-IL12 forms: Fcγ1- IL12 , Fcγ2- IL12, both forms, preferably with linker peptide and optionally modified Fc domain having reduced affinity to Fc-receptors.

EXAMPLE 13: *De-immunized forms of Fc-TNFa*

Non-deimmunized Fc- TNFa was described e.g. in WO 99/43713. Specific de-immunized Fc-TNFa forms: Fcγ1- TNFa , Fcγ2- TNFa , both forms, preferably with linker peptide and optionally modified Fc domain having reduced affinity to Fc-receptors.

EAMPLE 14: *De-immunized forms of Fc-GM-CSF*

Non-deimmunized Fc- GM-CSF was described e.g. in WO 99/43713 and WO 01/07081. Specific de-immunized Fc- GM-CSF forms: Fcγ1- GM-CSF , Fcγ2- GM-CSF , both forms, preferably with linker peptide and optionally modified Fc domain having reduced affinity to Fc-receptors.

EXAMPLE 15: *De-immunized forms of Fc-subtilisin*

Specific de-immunized Fc- subtilisin forms: Fcγ1- subtilisin , Fcγ2- subtilisin, both forms, preferably with linker peptide and optionally modified Fc domain having reduced affinity to Fc-receptors.

EXAMPLE 16: *De-immunized forms of Fc-insulin*

Specific de-immunized Fc- insulin forms: Fcγ1- insulin , Fcγ2- insulin, both forms, preferably with linker peptide and optionally modified Fc domain having reduced affinity to Fc-receptors.

EXAMPLE 17: *De-immunized forms of Fc-PSMA*

Non-deimmunized Fc- PSMA was described e.g. in WO 96/08570 and WO 01/0708. Specific de-immunized Fc- PSMA forms: Fcγ1- PSMA , Fcγ2- PSMA, both forms, preferably with linker peptide and optionally modified Fc domain having reduced affinity to Fc-receptors.

EXAMPLE 18:

De-immunized fusion proteins comprising anti-EGFR antibodies fused to a cytokine

Humanized and murine monoclonal antibody 425 (hMAb 425, US 5,558,864; EP 0531 472), murine and chimeric monoclonal antibody 225 (cMAb 225, US 4,943,533 and EP 0359 282) and murine and humanized MAb 4D5 (hMab 4D5 = Herceptin®) have been de-immunized according to the invention and fused to a de-immunized IL2 or a non-modified IL-2.

Fusions of antibodies to cytokines represent a situation where the need to reduce immunogenicity is particularly great. Normally, therapeutic antibodies can induce anti-idiotype antibodies that neutralize the effectiveness of a therapeutic antibody. This is particularly true when a therapeutic antibody is administered at low or medium levels, as opposed to very high levels where tolerance can be induced. For example, the therapeutic antibodies Herceptin and Rituxan are generally given in high doses of a few hundred milligrams. In contrast, antibody-cytokine fusions are generally given in a lower dose on the order of a few milligrams. Thus, the dose of an antibody-cytokine fusion is in the range that tends to promote formation of anti-idiotype antibodies. The presence of the linked cytokine tends to exaggerate the immunogenicity of the already immunogenic antibodies.

Antibody 425 is a non-human antibody which is directed to antigen EGF-R and reacts with colon cancer cells. This antibody has been fused to IL-2, as described in Example 13. The presence of IL-2 or another cytokine enhances the immunogenicity of the antibody, in particular the V regions.

In the following paragraphs the invention is described in more detail for the monoclonal anti-EGFR antibody 425- IL2 construct which was shown to have a high therapeutic value. However, the invention is not limited to this antibody and said construct and its several existing forms, but can be extended to other anti-EGFR antibodies and their fusion constructs,

preferably cytokine fusion immunoglobulins, above all chimeric antibody 225 (c225 – Il-2), which has very similar properties. In prinicipal, non-human, chimeric or humanized versions of the anti-EGFR antibodies can be used to synthesize said IL-2 fusion molecules

Unless stated otherwise all amino acids in the variable heavy and light chains are numbered as in Kabat et al, 1991 (Sequences of Proteins of Immunological Interest, US Department of Health and Human Services). Potential T-cell epitopes are numbered with the linear number of the first amino acid of an epitope, counting from the first amino acid of the heavy and light chains.

1. Comparison with Mouse Subgroup Frameworks

The amino acid sequences of murine 425 VH (heavy chain) and VK (light chain) were compared to consensus sequences for the Kabat murine heavy and light chain subgroups. 425 VH can be assigned to mouse heavy chains subgroup IIB. The comparison with the consensus sequence for this subgroup shows that the serine at position 94 in 425 VH is unusual. The most common residue at this position is arginine. 425 VK can be assigned to mouse kappa chains subgroup Vl. The comparison with the consensus sequence for this subgroup shows that the residues at positions 45-47, 60 and 100 in 425 VK are unusual for this subgroup. Amino acid residue numbering is as per Kabat.

2. Comparison with Human Frameworks

The amino acid sequences of murine 425 VH (variable heavy chain) and VK (variable kappa light chain) were compared to the sequences of the directory of human germline VH (Tomlinson, I.M et al., (1992) J. Mol.Biol. 227: 776-798) and VK (Cox, J.P.L. et al., (1994) Eur. J. Immunol. 24:827-836) sequences and also to human germline J region sequences (Routledge, E.G. et al., in, *Protein Engineering of Antibody Molecules for Prophylactic and Therapeutic Applications in Man*, Clark, M. ed. Academic Titles, Nottingham, UK, pp13-44, 1991). The murine 425 sequence of the heavy and light chain can be taken, for example, from EP 0531 472.

The reference human framework selected for 425 VH was VH1GRR with human JH6. The sequence of VH1GRR in the directory ends at residue 88. Therefore there is no corresponding residue for the unusual serine at position 94 of the murine sequence. This germline sequence has been found in a rearranged mature antibody gene with 4 amino acid changes. The reference human framework selected for 425 VK was L6/vg with human JK2. This germline sequence has been found in a rearranged mature antibody heavy chain with no amino acid changes.

3. Design of "veneered" sequences

Following identification of the reference human framework sequences, certain non-identical amino acid residues within the 425 VH and VK frameworks were changed to the corresponding amino acid in the human reference framework sequence. Residues which are considered to be critical for antibody structure and binding were excluded from this process and not altered. The murine residues that were retained at this stage are largely non-surface, buried residues, apart from residues at the N-terminus for instance, which are close to the CDRs in the final antibody (1 – 8, preferably 1 - 5 amino acid residues). This process produces a sequence that is broadly similar to a "veneered" antibody as the surface residues are mainly human and the buried residues are as in the original murine sequence.

4. Peptide Threading Analysis

The murine and veneered 425 VH and VK sequences were analyzed using the method according of the invention. The amino acid sequences are divided into all possible 13mers. The 13mer peptides are sequentially presented to the models of the binding groove of the HLA-DR allotypes and a binding score assigned to each peptide for each allele. A conformational score is calculated for each pocket-bound side chain of the peptide. This score is based on steric overlap, potential hydrogen bonds between peptide and residues in the binding groove, electrostatic interactions and favorable contacts between peptide and pocket residues. The conformation of each side chain is then altered and the score recalculated. Having determined the highest conformational score, the binding score is then calculated based on the groove-bound hydrophobic residues, the non-groove hydrophilic residues and the number of residues that fit into the binding groove. Peptides which are known binders to human MHC Class II achieve a high binding score with almost no false negatives. Thus peptides that achieve a significant binding score in the current analysis are considered to be potential T-cell epitopes. The results of the peptide threading analysis are shown in Table 1 for 425 VH and 425 VK. Potential T Cell epitopes are referred to by the linear number of the first residue of the 13mer.

Table 1: Potential T-cell epitopes in murine and veneered 425 sequences

Sequence	Number of potential T-cell epitopes	Number of first residue of 13mer with number of bonding alleles in brackets
Murine 425 VH	8	31(7), 35(17), 43(7), 46(8), 58(10), 62(12), 81(11), 84(16)
Veneered 425 VH	7	31(7), 43(7), 46(8), 58(10), 62(11), 81(11), 84(16)

Sequence	Number of potential T-cell epitopes	Number of first residue of 13mer with number of bonding alleles in brackets
Murine 425 VK	9	1(8), 2(5), 17(5), 27(5), 43(16), 72(18), 75(10), 92(10), 93(17)
Veneered 425 VK	4	27(5), 43(16), 92(8), 93(17)

5. Removal of Potential T Cell Epitopes

The numbering of amino acid residues for substitution is as per Kabat. Potential T Cell epitopes are referred to by the linear number of the first residue of the 13mer.

The amino acid substitutions required to remove the potential T-cell epitopes from the veneered 425 heavy chain variable region were as follows:

- Substitution of proline for alanine at residue 41 (Kabat number 41) removes the potential epitope at residue number 31.
- Substitution of proline for leucine at residue 45 (Kabat number 45) removes the potential epitope at residue number 43. A proline at position 45 is found in a human germline VH sequence, DP52.
- Substitution of alanine for isoleucine at residue 48 (Kabat number 48) removes the potential epitope at residue number 46.
- Substitution of valine for alanine at residue 68 (Kabat number 67) removes the potential epitope at residue number 58.
- Substitution of isoleucine for leucine at residue 70 (Kabat number 69) removes the potential epitope at residue number 62.
- Substitution of threonine for serine at residue 91 (Kabat number 87) removes the potential epitopes at residue numbers 81 and 84.

The amino acid substitutions required to remove the potential T-cell epitopes from the veneered 425 light chain variable region were as follows:

- Substitution of histidine for tyrosine at residue 35 (Kabat number 36) removes the potential epitope at residue number 27
- Substitution of alanine for threonine at residue 50 (Kabat number 51) removes the potential epitope at residue number 43. This residue is within CDR2. Alanine is commonly found at this position in both human and murine antibodies. An alternative substitution to eliminate this epitope is alanine for leucine at position 45 (Kabat number

46). There is no conservative substitution that will eliminate the potential epitope. Alanine is found at this position in some antibodies.

- Substitution of proline for isoleucine at residue 94 (Kabat number 95) removes the potential epitope at residue number 92. Kabat residue 95 is within CDRL3. Proline is common at this position in mouse antibody sequences and there is no change outwith the CDR that eliminates the potential epitope.
- Substitution of valine for leucine at residue 103 (Kabat number 104) removes the potential epitope at residue number 93.

6. Design of de-immunized Sequences

De-immunized heavy and light chain variable region sequences were designed with reference to the changes required to remove potential T-cell epitopes and consideration of framework residues that might be critical for antibody structure and binding. In addition to the De-immunized sequences based on the veneered sequence. an additional sequence was designed for each of VH and VK based on the murine sequence, termed the Mouse Peptide Threaded (Mo PT) version. For this version, changes were made directly to the murine sequence in order to eliminate T-cell epitopes, but only changes out with the CDRs that are not considered to be detrimental to binding are made. No attempt to remove surface (B-cell) epitopes has been made in this version of the de-immunized sequence.

The primary de-immunized VH includes substitutions 1 to 6 in Section 5 above and includes no potential T-cell epitopes. A further 4 de-immunized VH sequences were designed in order to test the effect of the various substitutions required on antibody binding. The cumulative alterations made to the primary de-immunized sequence (425 VH1GRR-VH-v1) and the potential T-cell epitopes remaining are detailed in Table 2. The mouse threaded version is included for comparison.

Table 2: Amino acid changes and potential epitopes in de-immunized 425 VH

Variant	Cumulative Residue Changes	Potential T Cell Epitopes
425 VH1GRR-VH-v1	None	None
425 VH1GRR-VH-v2	48A → I	46(8)
425 VH1GRR-VH-v3	45P → L	43(7), 46(8)
425 VH1GRR-VH-v4	67V → A, 69I → L	43(7). 46(8), 58(10), 62(11)
425 VH1GRR-VH-v5	41P → A	31(7), 43(7), 46(8), 58(10), 62(11)
425 VH-MoPT	NA	43(7), 46(8)

The primary de-immunized VK includes substitutions 1 to 4 in Section 5 above and includes no potential T-cell epitopes. A further 4 de-immunized VK sequences were designed in order to test the effect of the various substitutions required on antibody binding. Version 2 is an alternative to Version 1 in which an alternative substitution has been used to remove the same potential T-cell epitope. The cumulative alterations made to the primary de-immunized sequence (425 L6-vg-VK-v1) and the potential T-cell epitopes remaining are detailed in Table 3. The mouse threaded version is included for comparison.

Table 3: Amino acid changes and potential epitopes in de-immunized 425 VK

Variant	Cumulative Residue Changes	Potential T-cell Epitopes
425 L6-vg-VK-v1	None	None
425 L6-vg-VK-v1	51 A → T, 46L → A	None
425 L6-vg-VK-v1	46 A → L	43(16)
425 L6-vg-VK-v1	95 P → I	43(16), 92(8)
425 L6-vg-VK-v1	36 H → Y	27(5), 43(16), 92(8)
425VK-MoPT	NA	27(5), 43(16), 92(8)

Table 4: *original and "veneered" sequences of VH and VK of murine MAb 425*

425 VH mouse

QVQLQQPGAELVKPGASVKLSCKASGYTFTSHWMHWVKQRAGQGLEWIGEFNPSNGRTNYNEKFKSKAT
LTVDKSSSTAYMQLSSLTSEDSAVYYCASRDYDYDGRYFDYWGQGTTLTVSS

425 VK mouse

QIVLTQSPAIMSASPGEKVTMTCSASSSVTYMYWYQQKPGSSPRLLIYDTSNLASGVPVRFSGSGSGTS
YSLTISRMEAEDAATYYCQQWSSHIFTFGSGTKLEIK

425 VH veneered:

QVQLVQSGAELVKPGASVKLSCKASGYTFTSHWMHWVKQAAGQGLEWIGEFNPSNGRTNYNEKFKSRAT
LTVDKSTSTAYMQLSSLTSEDSAVYYCASRDYDYDGRYFDYWGQGTTLTVSS

425 VK veneered:

QIVLTQSPATLSASPGERATMSCSASSSVTYMYWYQQKPGQSPRLLIYDTSNLASGVPARFSGSGSGTS
YTLTISSLEAEDAATYYCQQWSSHIFTFGQGTKLEIK

Table 5: *De-immunized sequences of variable heavy and light chain of MAb 425*

425 de-immunized VH1

QVQLVQSGAELVKPGASVKLSCKASGYTFTSHWMHWVKQAPGQGPEWAGEFNPSNGRTNYNEKFKSRVT
ITVDKSTSTAYMQLSSLTSEDTAVYYCASRDYDYDGRYFDYWGQGTTLTVSS

425 de-immunized VK1

QIVLTQSPATLSASPGERATMSCSASSSVTYMYWHQQKPGQSPRLLIYDASNLASGVPARFSGSGSGTS
YTLTISSLEAEDAATYYCQQWSSHPFTFGQGTKVEIK

425 de-immunized VH2

QVQLVQSGAELVKPGASVKLSCKASGYTFTSHWMHWVKQAPGQGPEWIGEFNPSNGRTNYNEKFKSRVT
ITVDKSTSTAYMQLSSLTSEDTAVYYCASRDYDYDGRYFDYWGQGTTLTVSS

425 de-immunized VK2

QIVLTQSPATLSASPGERATMSCSASSSVTYMYWHQQKPGQSPRALIYDTSNLASGVPARFSGSGSGTS
YTLTISSLEAEDAATYYCQQWSSHPFTFGQGTKVEIK

425 de-immunized VH3

QVQLVQSGAELVKPGASVKLSCKASGYTFTSHWMHWVKQAPGQGLEWIGEFNPSNGRTNYNEKFKSRVT
ITVDKSTSTAYMQLSSLTSEDTAVYYCASRDYDYDGRYFDYWGQGTTLTVSS

425 de-immunized VK3

QIVLTQSPATLSASPGERATMSCSASSSVTYMYWHQQKPGQSPRLLIYDTSNLASGVPARFSGSGSGTS
YTLTISSLEAEDAATYYCQQWSSHPFTFGQGTKVEIK

425 de-immunized VH4

QVQLVQSGAELVKPGASVKLSCKASGYTFTSHWMHWVKQAPGQGLEWIGEFNPSNGRTNYNEKFKSRAT
LTVDKSTSTAYMQLSSLTSEDTAVYYCASRDYDYDGRYFDYWGQGTTLTVSS

425 de-immunized VK4

QIVLTQSPATLSASPGERATMSCSASSSVTYMYWHQQKPGQSPRLLIYDTSNLASGVPARFSGSGSGTS
YTLTISSLEAEDAATYYCQQWSSHIFTFGQGTKVEIK

425 de-immunized VH5

QVQLVQSGAELVKPGASVKLSCKASGYTFTSHWMHWVKQAAGQGLEWIGEFNPSNGRTNYNEKFKSRAT
LTVDKSTSTAYMQLSSLTSEDTAVYYCASRDYDYDGRYFDYWGQGTTLTVSS

425 de-immunized VK5

QIVLTQSPATLSASPGERATMSCSASSSVTYMYWYQQKPGQSPRLLIYDTSNLASGVPARFSGSGSGTS
YTLTISSLEAEDAATYYCQQWSSHIFTFGQGTKVEIK

425 VH mouse, peptide threaded (Mo PT)

QVQLQQPGAELVKPGASVKLSCKASGYTFTSHWMHWVKQAPGQGLEWIGEFNPSNGRTNYNEKFKSRVT
ITVDKSSSTAYMQLSSLTSEDTAVYYCASRDYDYDGRYFDYWGQGTTLTVSS

425 VK mouse, peptide threaded (Mo PT)

QIVLTQSPATLSASPGEKATMTCSASSSVTYMYWYQQKPGSSPRLLIYDTSNLASGVPVRFSGSGSGTS
YSLTISRLEAEDAATYYCQQWSSHIFTFGQGTKVEIK

As already mentioned, the modified anti-EGFR antibody – cytokine constructs according to the invention, preferably MAb 425 – Il2, can be used in pharmaceutical compositions and pharmaceutical kits preferably for the treatment of cancer. "Cancer" and "tumor" refer to or

describe the physiological condition in mammals that is typically characterized by unregulated cell growth. By means of the pharmaceutical compositions according of the present invention tumors can be treated such as tumors of the breast, heart, lung, small intestine, colon, spleen, kidney, bladder, head and neck, ovary, prostate, brain, pancreas, skin, bone, bone marrow, blood, thymus, uterus, testicles, cervix, and liver.

In analogy to antibody 425 similar fusion constructs can be obtained using monoclaonal antibody 225 in murine, chimeric or humanized forms.

EXAMPLE 19: *De-immunized forms of 14.18 antibody – IL2 and KS- 1/4 – IL2 .*

The cytokine interleukin 2 (IL-2) has been fused to specific monoclonal antibodies KS-1/4 and ch14.18 directed to the tumor associated antigens epithelial cell adhesion molecule (Ep-CAM, KSA, KS1/4 antigen) and the disialoganglioside GD, respectively, to form the fusion proteins ch14.18-IL-2 and KS1/4-IL-2, respectively, (US 5,650,150, EP 0 338 767). Theses antibodies have been de-immunized according to the invention and fused to a immunogenicly modified IL2 or a non-modified IL-2.

Anti-EpCAM antibody KS 1/4

The monoclonal antibody KS1/4 is a murine antibody that specifically binds to the 40,000 dalton cell surface antigen EpCAM (epithelial cell adhesion molecule) found in high density on adenocarcinoma cells and also found at much lower levels on certain normal epithelial cells. This antibody has been shown to be effective for the detection of disease.

A variety of fusions of KS-1/4 to single and combined cytokines such as IL-2 and IL-12, have been described (WO98/25978, WO01/58957A, and WO 01/10912). These fusion proteins are effective in animal models of cancer. However, due to the presence of cytokines, these fusion proteins are particularly immunogenic. There is a need for altered KS antibody molecules with a reduced propensity to elicit an immune response on administration to the human host. Modified sequences in Mab KS1/4 providing a modified KS antibody according to the invention are shown below. A mutated form of the KS-1/4 in which the T-cell epitopes in the V regions were completely removed by mutation, as defined by the criteria given above in the section on computer algorithms, was efficiently expressed in mammalian cells and bound to the EpCAM antigen with only about an 8-fold reduction of affinity. This molecule was termed VHv1/VKv1. A second antibody, VHv2/VKv1, had only about a 3-fold reduction in affinity and differed from VHv1/VKv1 by a single amino acid substitution. These two antibodies have been expressed in mammalian cells as KS-IL2 fusion proteins. The

KS(VHv1/VKv1)-IL2 and KS(VHv2/VKv1) are the most preferred embodiments of the invention with respect to treatment of a broad spectrum of human cancers by immune therapy.

1 Comparison with Mouse Subgroup Frameworks

The amino acid sequences of murine KS VH and VK were compared to consensus sequences for the Kabat murine heavy and light chain subgroups (Kabat et al., 1991). Murine KS VH cannot be assigned to any one Subgroup, but is closest to Subgroup II(A) and V(A). Unusual residues are found at position 2 which is normally valine, 46 which is normally glutamic acid, and 68 which is normally threonine. Residue 69 is more commonly leucine or iso-leucine. At 82b, serine is most often found. Murine KS VK can be assigned to Subgroup VI (' Figure 2). Unusual residues are found at 46 and 47 which are commonly both leucine. Residue 58 is unusual with either leucine or valine normally found at this position.

2 Comparison with Human Frameworks

The amino acid sequences of murine KS VH and VK were compared to the sequences of the directory of human germline VH (Tomlinson et al., 1992) and VK (COX et al. 1994) sequences and also to human germline J region sequences (Routledge et al., 1993). The reference human framework selected for KS VH was DP10 with human JH6. This germline sequence has been found in a rearranged mature antibody gene with no amino acid changes. The reference human framework selected for KS VK was B1. For framework- 2 the sequence of the mature human antibody IMEV was used (in Kabat et al 1991). This sequence is identical to the murine sequence immediately adjacent to CDR2. The J region sequence was human JK4. This germline sequence has not been found as rearranged mature antibody light chain.

3 Design of Veneered Sequences

Following identification of the reference human framework sequences, certain non-identical amino acid residues within the 425 VH and VK frameworks were changed to the corresponding amino acid in the human reference sequence. Residues which are considered to be critical for antibody structure and bindin2 were excluded from this process and not altered. The murine residues that were retained at this stage are largely non-surface, buried residues, apart from residues at the N-terminus for instance, which are close to the CDRs in the final antibody. This process produces a sequence that is broadly similar to a "veneered" antibody as the surface residues are mainly human and the buried residues are as in the original murine sequence.

4 Peptide Threading Analysis

The murine and veneered KS VH and VK sequences were analyzed using the method according to the invention. The amino acid sequences are divided into all possible 13imers. The 13-mer peptides are sequentially presented to the models of the binding groove of the HLA-DR allotypes and a binding score assigned to each peptide for each allele. A conformational score is calculated for each pocket-bound side chain of the peptide. This score is based on steric overlap, potential hydrogen bonds between peptide and residues in the binding groove, electrostatic interactions and favorable contacts between peptide and pocket residues. The conformation of each side chain is then altered and the score recalculated. Having determined the highest conformational score, the binding score is then calculated based on the (groove-bound hydrophobic residues, the non-groove hydrophilic residues and the number of residues that fit into the binding groove. Known binders to MHC class II achieve a significant binding score with almost no false negatives. Thus peptides achieving, a significant binding score from the current analysis are considered to be potential T-cell epitopes. The results of the peptide threading analysis for the murine and veneered sequences are shown in Table 1.

Table 1: Potential T-cell epitopes in murine and veneered KS sequences

Sequence	Number of potential T-cell epitopes	Location of potential epitopes (no. of potential MHC binders)
Murine KS VH	6	35(11), 62(17), 78(12), 81(12), 89(6), 98(15)
Murine KS VH	5	30(7), 62(15), 78(11), 89(6), 98(15)
Murine KS VK	6	1(14), 2(5), 17(5), 27(5), 51(13), 72(18)
Veneered KS VK	3	1(17), 27(5), 51(13)

5 Removal of Potential T Cell Epitopes

Potential T-cell epitopes are removed by making amino acid substitutions in the particular peptide that constitutes the epitope. Substitutions were made by inserting amino acids of similar physicochemical properties if possible. However in order to remove some potential epitopes, amino acids of different size, charge or hydrophobicity may need to be substituted. If changes have to made within CDRs which might have an effect on binding, there is then a need to make a variant with and without the particular amino acid substitution. Numbering of amino acid residues for substitution is as per Kabat. Potential T Cell epitopes are referred to by the linear number of the first residue of the 13mer.

The amino acid changes required to remove T-cell epitopes from the veneered KS heavy chain variable region were as follows:

1. Substitution of arginine for lysine at residue 38 (Kabat number 38) removes the potential epitope at residue no 30.

2. Substitution of alanine for leucine at residue 72 (Kabat number 71) and isoleucine for phenylalanine at residue 70 (Kabat number 69) removes the potential epitope at residue 62. An isoleucine at Kabat number 69 and alanine at Kabat number 71 is found in a human germline VH sequence, DP10.

3. Substitution of leucine for alanine at residue 79 (Kabat number 78) removes the potential epitope at residue number 78.

4. Substitution of threonine for methionine at residue 91 (Kabat number 87), removes the potential epitope at residue number 89.

5. Substitution of methionine for at isoleucine residue 100 (Kabat number 96) in CDRH3 removes the potential epitope at residue 98. There is no change out with CDRH3 which removes this potential epitope.

The amino acid substitutions required to remove the potential T-cell epitopes from the veneered KS light chain variable region were as follows:

1. Substitution of isoleucine for methionine at residue 32 (Kabat number 33) removes the potential epitope at residue number 27. This residue is within CDR2. Isoleucine is commonly found at this position in human antibodies.

2. The potential epitope at position 1 is removed by substituting valine for leucine at residue (Kabat number 3).

3. Substitution of serine for alanine at residue 59 (Kabat number 60) removes the potential epitope at residue number 51.

6 Design of de-immunized Sequences

De-immunized heavy and light chain sequences were designed with reference to the changes required to remove potential T-cell epitopes and consideration of framework residues that might be critical for antibody structure and binding. In addition to the de-immunized sequences based on the veneered sequence, an additional sequence was designed for each VH, and VK based on the murine sequence, termed the Mouse Peptide Threaded (MoPT) version. For this version, changes. were made directly to the murine sequence in order to eliminate T-cell epitopes, but only changes outside the CDRs that are not considered to be detrimental to binding are made. No attempt to remove surface (B cell) epitopes has been made in this

version of the de-immunized sequence. The primary de-immunized VH includes substitutions 1 to 5 in Section 5 above and one extra change at residue 43 (Kabat number 43). Lysine found in the murine sequence was substituted for the glutamine from the human framework. Lysine is positively charged and therefore significantly different to glutamine; this region may be involved in VH/VL contacts. The primary de-immunized VH includes no potential T-cell epitopes. A further 4 de-immunized VHs were designed in order to test the effect of the various substitutions required on antibody binding. The cumulative alterations made to the primary de-immunized sequence (KSDIVHv1) and the potential T-cell epitopes remaining are detailed in Table 2.

Table 2: Amino acid changes and potential epitopes in de-immunized KS VH

Variant	Cumulative residue changes	Potential epitopes (no. of potential MHC binders from 18 tested)
KSDIVHv1	None	none
KSDIVHv2	96M → I	98(15)
KSDIVHv3	71A → L, 78L → A	62(16), 78(11), 98(15)
KSDIVHv4	38 R → K	30(7), 62(16), 78(11), 98(15)
KSDIVHv5	68T → A, 69I → F	30(7), 62(17), 78(11), 98(15)
KSMoPTVH	NA	98(15), 78(12)

The primary de-immunized VK includes substitutions 1 to 3 in Section 5 above. A further 3 de-immunized VKs were designed in order to test the effect of the various substitutions required on antibody binding. The cumulative alterations made to the primary de-immunized sequence (KSDIVKv1) and the potential T-cell epitopes remaining are detailed in Table 3.

Table 3: Amino acid changes and potential epitopes in de-immunized KS VK

Variant	Cumulative residue changes	Potential epitopes (no. of potential MHC binders from 18 tested)
KSDIVKv1	None	none
KSDIVKv2	33I → M	27(5)
KSDIVKv3	3V → L	1(17), 27(5)
KSDIVKv4	60 S → A	1(17), 27(5), 5(13)
KSMoPTVK	NA	none

Sequences of versions of modified epitopes:

KS VH veneered:

QIQLVQSGPELKKPGSSVKiSCKASGYTFTNYGMNWVKQAPGQGLKWMGWINTYTGEPTYADDFKGRFT
fTlETSTSTAYLQLNNLRsEDmATYfCVRFISKGDYWGQGTTVTVSS

KS VK veneered:

QILLTQSPASLAVSPGQRATITCSASSSVSYMLWYQQKPGQPPKPWIFDTSNLASGFPARFSGSGSGTS
YTLTINSLEAEDAATYYCHQRSGYPYTFGGGTKVEIK

KS de-immunized VH1

QIQLVQSGPELKKPGSSVKISCKASGYTFTNYGMNWVRQAPGKGLKWMGWINTYTGEPTYADDFKGRFT
ITAETSTSTLYLQLNNLRSEDTATYFCVRFMSKGDYWGQGTTVTVSS

KS de-immunized VK1

QIVLTQSPASLAVSPGQRATITCSASSSVSYILWYQQKPGQPPKPWIFDTSNLASGFPSRFSGSGSGTS
YTLTINSLEAEDAATYYCHQRSGYPYTFGGGTKVEIK

KS de-immunized VH2

QIQLVQSGPELKKPGSSVKISCKASGYTFTNYGMNWVRQAPGKGLKWMGWINTYTGEPTYADDFKGRFT
ITAETSTSTLYLQLNNLRSEDTATYFCVRFISKGDYWGQGTTVTVSS

KS de-immunized VK2

QIVLTQSPASLAVSPGQRATITCSASSSVSYMLWYQQKPGQPPKPWIFDTSNLASGFPSRFSGSGSGTS
YTLTINSLEAEDAATYYCHQRSGYPYTFGGGTKVEIK

KS de-immunized VH3

QIQLVQSGPELKKPGSSVKISCKASGYTFTNYGMNWVRQAPGKGLKWMGWINTYTGEPTYADDFKGRFT
ITLETSTSTAYLQLNNLRSEDTATYFCVRFISKGDYWGQGTTVTVSS

KS de-immunized VK3

QILLTQSPASLAVSPGQRATITCSASSSVSYMLWYQQKPGQPPKPWIFDTSNLASGFPSRFSGSGSGTS
YTLTINSLEAEDAATYYCHQRSGYPYTFGGGTKVEIK

KS de-immunized VH4

QIQLVQSGPELKKPGSSVKISCKASGYTFTNYGMNWVKQAPGKGLKWMGWINTYTGEPTYADDFKGRFT
ITLETSTSTAYLQLNNLRSEDTATYFCVRFISKGDYWGQGTTVTVSS

KS de-immunized VK4

QILLTQSPASLAVSPGQRATITCSASSSVSYMLWYQQKPGQPPKPWIFDTSNLASGFPARFSGSGSGTS
YTLTINSLEAEDAATYYCHQRSGYPYTFGGGTKVEIK

KS de-immunized VH5

QIQLVQSGPELKKPGSSVKISCKASGYTFTNYGMNWVKQAPGKGLKWMGWINTYTGEPTYADDFKGRFA
FTLETSTSTAYLQLNNLRSEDTATYFCVRFISKGDYWGQGTTVTVSS

KS de-immunized VK5

QILLTQSPASLAVSPGQRATITCSASSSVSYMLWYQQKPGSSPKPWIYDTSNLASGFPARFSGSGSGTS
YTLTINSLEAEDAATYYCHQRSGYPYTFGGGTKVEIK

KS VH mouse, peptide threaded (Mo PT)

QIQLVQSGPELKKPGETVKISCKASGYTFTNYGMNWVRQAPGKGLKWMGWINTYTGEPTYADDFKGRFV
FSLETSASTAFLQLNNLRSEDTATYFCVRFISKGDYWGQGTSVTVSS

KS VK mouse, peptide threaded (Mo PT)

QIVLTQSPATLSASPGERVTITCSASSSVSYMLWYLQKPGSSPKPWIFDTSNLASGFPSRFSGSGSGTT
YSLIISSLEAEDAATYYCHQRSGYPYTFGGGTKLEIK

KS VH mouse

QIQLVQSGPELKKPGETVKISCKASGYTFTNYGMNWVKQTPGKGLKWMGWINTYTGEPTYADDFKGRFA
FSLETSASTAFLQINNLRNEDMATYFCVRFISKGDYWGQGTSVTVSS

KS VK mouse

QILLTQSPAIMSASPGEKVTMTCSASSSVSYMLWYQQKPGSSPKPWIFDTSNLASGFPARFSGSGSGTS
YSLIISSMEAEDAATYYCHQRSGYPYTFGGGTKLEIK

MAb 14.18 – IL-2

In analogy, monoclonal antibody 14.18 was fused to IL-2 and deimmunized according to the invention.

Potential T-cell epitopes in murine and veneered 14.18 sequences are:

Sequence	Number of potential T-cell	Location of potential epitopes
Murine 14,18 VH	11	3(17), 9(15), 30(5), 35(17), 39(15), 43(9), 58(12), 62(11), 81(11), 84(16), 101(7)
Veneered 14. 18 VH	5	43(9), 58(12), 62(11), 81(11), 84(16)
Murine 14.18 VK	7	7(7), 13(11), 27(15), 49(11), 86(17), 97(11), 100(4)
Veneered 14. 18 VK	5	27(I5), 49(11), 86(17), 97(11), 100(17)

Amino acid changes and potential epitopes in de-immunized 14.18 VH are:

Variant	Cumulative residue changes	Potential epitopes (no. of potential MHC binders from 18 tested)
14.18DIVH1	none	none
14.18DIVH2	41I → P, 45L → T, 50L → A	none

14.18DIVH3	65S → G	58(8)
14.18DIVH4	71A → V	58(8), 62(4)
14.18DIVH5	45T → L, 41P → 1	43(9) 58(8) 62(4)
14.18MoPTVH	NA	43(9) 58(12) 62(11)

Amino acid changes and potential epitopes in de-immunized 14.18 VK are:

Variant	Cumulative residue changes*	Potential epitopes' (no. of potential MHC binders from 18 tested)
14.18DIVKI	None	none
14.18DIVK2	46L→ M, 49Y → H	none
14.18DIVK3	96P → T, 100Q → G	97(5)
14.18DIVK4	96T → L	97(11)
14.18DIVK5	27e S → R	27(15), 97(11)
14.18DIVK6	46M → L	27(15), 49 (11), 97(11)
14.18MoPTVK	NA	27(15), 49 (11), 97(11), 100(4)

Sequences of versions of modified epitopes are:

14.18 VH veneered:

EVQLLQSGPELKKPGASVKISCKASGSSFTGYNMNWVRQAPGQRLEWIGAIDPYYGGTSYNQKFKGRAT
LSVDKSSSQAYMHLKSLTSEDSAVYYCVSGMEYWGQGTTVTVSS

14.18 VK veneered:

DVVMTQSPGTLPVSLGERATISCRSSQSLVHRNGNTYLHWYLQKPGQSPKLLIHKVSNRFSGVPDRFSG
SGSGTDFTLTISRLEAEDLAVYFCSQSTHVPPLTFGQGTKLEIK

14.18 de-immunized VH1

EVQLLQSGPELKKPGASVKISCKASGSSFTGYNMNWVRQAIGQRLEWIGLIDPYYGGTSYNQKFKSRVT
ITADKSSSQAYMHLKSLTSEDTAVYYCVSGMEYWGQGTTVTVSS

14.18 de-immunized VK1

DVVMTQSPGTLPVSLGERATISCRSSQSLVHSNGNTYLHWYLQKPGQSPKLLIYKVSNRFSGVPDRFSG
SGSGTDFTLTISRLEAEDMAVYFCSQSTHVPPPTFGQGTKVEIK

14.18 de-immunized VH2

EVQLLQSGPELKKPGASVKISCKASGSSFTGYNMNWVRQAPGQRTEWIGAIDPYYGGTSYNQKFKSRVT
ITADKSSSQAYMHLKSLTSEDTAVYYCVSGMEYWGQGTTVTVSS

14.18 de-immunized VK2

DVVMTQSPGTLPVSLGERATISCRSSQSLVHSNGNTYLHWYLQKPGQSPKMLIHKVSNRFSGVPDRFSG
SGSGTDFTLTISRLEAEDMAVYFCSQSTHVPPPTFGQGTKVEIK

14.18 de-immunized VH3

EVQLLQSGPELKKPGASVKISCKASGSSFTGYNMNWVRQAPGQRTEWIGAIDPYYGGTSYNQKFKGRVT
ITADKSSSQAYMHLKSLTSEDTAVYYCVSGMEYWGQGTTVTVSS

14.18 de-immunized VK3

DVVMTQSPGTLPVSLGERATISCRSSQSLVHSNGNTYLHWYLQKPGQSPKMLIHKVSNRFSGVPDRFSG
SGSGTDFTLTISRLEAEDMAVYFCSQSTHVPPTTFGGGTKVEIK

14.18 de-immunized VH4

EVQLLQSGPELKKPGASVKISCKASGSSFTGYNMNWVRQAPGQRTEWIGAIDPYYGGTSYNQKFKGRVT
ITVDKSSSQAYMHLKSLTSEDTAVYYCVSGMEYWGQGTTVTVSS

14.18 de-immunized VK4

DVVMTQSPGTLPVSLGERATISCRSSQSLVHSNGNTYLHWYLQKPGQSPKMLIHKVSNRFSGVPDRFSG
SGSGTDFTLTISRLEAEDMAVYFCSQSTHVPPLTFGGGTKVEIK

14.18 de-immunized VH5

EVQLLQSGPELKKPGASVKISCKASGSSFTGYNMNWVRQAIGQRLEWIGAIDPYYGGTSYNQKFKGRVT
ITVDKSSSQAYMHLKSLTSEDTAVYYCVSGMEYWGQGTTVTVSS

14.18 de-immunized VK5

DVVMTQSPGTLPVSLGERATISCRSSQSLVHRNGNTYLHWYLQKPGQSPKMLIHKVSNRFSGVPDRFSG
SGSGTDFTLTISRLEAEDMAVYFCSQSTHVPPLTFGGGTKVEIK

14.18 VH mouse, peptide threaded (Mo PT)

EVQLVQSGPEVEKPSASVKISCKASGSSFTGYNMNWVRQAIGKSLEWIGAIDPYYGGTSYNQKFKGRAT
LTVDKSSSTAYMHLKSLTSEDTAVYYCVSGMEYWGQGTTVTVSS

14.18 VK mouse, peptide threaded (Mo PT)

DVVMTQTPGSLPVSAGDQASISCRSSQSLVHRNGNTYLHWYLQKPGQSPKLLIHKVSNRFSGVPDRFSG
SGSGTDFTLKISRVEAEDSGVYFCSQSTHVPPLTFGAGTKLELK

14.18 VH mouse

EVQLLQSGPELEKPSASVMISCKASGSSFTGYNMNWVRQNIGKSLEWIGAIDPYYGGTSYNQKFKGRAT
LTVDKSSSTAYMHLKSLTSEDSAVYYCVSGMEYWGQGTSVTVSS

14.18 VK mouse

DVVMTQTPLSLPVSLGDQASISCRSSQSLVHRNGNTYLHWYLQKPGQSPKLLIHKVSNRFSGVPDRFSG
SGSGTDFTLKISRVEAEDLGVYFCSQSTHVPPLTFGAGTKLELK

The foregoing description and the examples are intended as illustrative, and are not to be taken as limiting. Still other variants within the spirit and scope of this invention are possible and will readily present themselves to those skilled in the art.

PATENT CLAIMS:

1. An immunogenicly modified fusion protein derived from a parent fusion protein, essentially consisting of a first protein / polypeptide and a second protein / polypeptide, wherein the first protein is an immunoglobulin molecule or a fragment thereof and the second protein / polypetide is non-immunoglobulin target polypeptide (X) each linked to the other directly or by a linker molecule; said modified fusion protein having an amino acid sequence different from that of said parent fusion protein and exhibiting reduced immunogenicity by a reduced number of T-cell epitopes within its amino acid sequence relative to the parent fusion protein when exposed to the immune system of a given species.

2. A modified fusion protein according to claim 1, wherein said T-cell epitopes are peptide sequences able to bind to MCH class II molecule binding groups.

3. A modified fusion protein of claim 1 or 2, wherein the target polypeptide (X) is linked by its N-terminal to the C-terminal of the immunoglobulin moiety.

4. A modified fusion protein according to any of the claims 1 – 3, wherein the given species is a human.

5. A modified fusion protein according to any of the claims 1 – 3, wherein the fusion components are fused via a linker molecule L.

6. A modified fusion protein according to claim 4, wherein said linker molecule L is non-immunogenic or less immunogenic.

7. A modified fusion protein according to any of the claims 1 – 6, wherein the fusion region represented the C-terminal region of the immunoglobulin portion and the N-terminal region of the non-immunoglobulin target polypeptide (X) has no a reduced number of T-cell epitopes.

8. A modified fusion protein according to any of the claims 1 - 7, wherein the immunoglobulin portion or a fragment thereof is less immunogenic.

9. A modified fusion protein according to any of the claims 1 - 7, wherein the target polypeptide (X) portion is less immunogenic.

10. A modified fusion protein according to any of the claims 1 – 9, wherein said immunoglobulin molecule or fragment thereof is IgG1 or IgG2.

11. A modified fusion protein according to any of the claims 1 – 10 , wherein said immunoglobulin fragment is a Fc portion.

12. A modified fusion protein according to claim 11, wherein said Fc portion has a reduced affinity to Fc receptors.

13. A modified fusion protein according to claim 11 or 12 having the formula

$$Fc - L_n - X$$

wherein

Fc is the Fc portion of an immunoglobulin molecule (antibody),

X is a non-immunoglobulin target polypeptide

L is a linker peptide,

n = 0 or 1, and

wherein X and / or L comprises amino acid residue modifications which elicit a reduced immunogenicity compared to the parent molecule.

14. A modified fusion protein according to claim 13, wherein at least X has no a reduced immunogenicity.

15. A modified molecule according to claim 14, wherein furthermore the fusion region between Fc and X and optionally Fc and L and / or L and X has no or a reduced number of T-cell epitopes.

16. A modified fusion protein according to any of the claims 1 – 10 having the formula

$$A - L_n - X$$

wherein

A is a whole antibody or its sFv, Fab, Fab', $F(ab')_2$ fragments

X is a non-immunoglobulin target polypeptide

L is a linker peptide,

n = 0 or 1, and

wherein A and / or X and / or L comprise amino acid residue modifications which elicit a reduced immunogenicity compared to the parent molecule.

17. A modified fusion protein according to claim 16, wherein at least A or X has no or a reduced number of T-cell epitopes.

18. A modified molecule according to claim 17, wherein furthermore the fusion region between A and X and optionally A and L and / or L and X has no or a reduced immunogenicity.

19. A modified fusion protein according to any of the claims 16 – 18, wherein A is selected from the group:

anti- EGF receptor (HER1) antibodies

anti- HER2 antibodies

anti- CDx antibodies

anti- cytokine receptor antibodies

anti- 17-1A antibodies,

anti- KSA antibodies

anti-GP IIb/IIIa antibodies

anti-integrin receptor antibodies

anti VEGF receptor antibodies.

20. A modified fusion protein according to claim 19, wherein the antibody is selected from the group:

monoclonal antibody 225 and derivatives,

monoclonal antibody 425 and derivatives

monoclonal antibody KS 1/4 and derivatives
monoclonal antibody 14.18 and derivatives
monoclonal antibody 4D5 / HER2 (Herceptin®) and derivatives
monoclonal antibody 17-1A and derivatives
monoclonal antibody 7E3 and derivatives
monoclonal antibodies LM609, P1F6 and 14D9.F8 and derivatives
monoclonal antibody DC-101 and derivatives
monoclonal anti-Il-2R antibody (Zenapax®) and derivatives

21. A modified fusion protein according to any of the claims 1 – 20, wherein the target polypeptide X is selected from the group:
cytokines, integrin inhibitors, soluble cytokine receptors, glycoproteins, hormones, glycoprotein hormones, leptin, growth hormones, growth factors, anti-hemophilic factors, antigens, cytokine receptor antagonists.

22. A modified fusion protein according to claim 21, wherein the target polypeptide X is selected from the group:
IL-2, G-CSF, GM-CSF, EPO, TPO, TNFα, soluble TNF receptor, IL-12, IL-8, FGF, TGF, EGF, VEGF, PMSA, IGF, insulin, hGH, RGD-peptides, endostatin, angiostatin, BDNF, CNTF, protein c, factor IX, and
and biologically active fragments thereof.

23. A modified fusion protein according to any of the claims 1 – 22, selected from the group:
MAb KS 1/4 – IL2, MAb 14.18 – IL2
MAb 425 – IL2, MAb c425 – IL2, MAb h425 – IL2, MAb 425 – TNFa
MAb 225 – IL2, MAb c225 – IL2
MAb 4D5 – IL2, MAb DC101 – Il2, MAb LM609 – IL2,
Fc – IL2 , Fc – TNFa, Fc – G-CSF, Fc – EPO, Fc – Leptin, Fc – KGF,
Fc – BDNF, Fc-CNTF, FC – ß-Cerebrosidase, Fc – TPO, Fc – GM-CSF,

24. A DNA sequence encoding a fusion protein of any of the claims 1 – 23 and 36.

25. A DNA sequence of claim 24, comprising
(i) a signal sequence
(ii) a DNA sequence encoding all domains or a Fc, sFV, Fab, Fab' or F(ab')2 domain of an IgG1, IgG2 or IgG3 antibody, and
(ii) a DNA sequence encoding the polypeptide (X), and optionally
(iii) a DNA sequence encoding the linker molecule.

26. An expression vector comprising a DNA sequence of claims 24 or 25.

27. A pharmaceutical composition comprising a fusion protein according any of the claims 1 – 23 and 36, optionally together with a suitable carrier, excipient or diluent.

28. A method for preparing an immunogenicly modified fusion protein as specified in of the claims 1 – 23 comprising the steps:
(i) determining the amino acid sequence of the parent fusion protein or part thereof;
(ii) identifying one or more potential T-cell epitopes within the amino acid sequence of the fusion protein by any method including determination of the binding of the peptides to MHC molecules using *in vitro* or *in silico* techniques or biological assays, (iii) designing new sequence variants by alteration of at least one amino acid residue within the originally identified T-cell epitope sequences, said variants are modified in such a way to substantially reduce or eliminate the activity or number of the T-cell epitope sequences and / or the number of MHC allotypes able to bind peptides derived from said biological molecule as determined by the binding of the peptides to MHC molecules using *in vitro* or *in silico* techniques or biological assays or by binding of peptide-MHC complexes to T-cells, (iv) constructing such sequence variants by recombinant DNA techniques and testing said variants in order to identify one or more variants with desirable properties, and (v) optionally repeating steps (ii) – (iv),
characterized in that the identification of T-cell epitope sequences according to step (ii) is achieved by
(a) selecting a region of the peptide having a known amino acid residue sequence;
(b) sequentially sampling overlapping amino acid residue segments of predetermined uniform size and constituted by at least three amino acid residues from the selected region; (c) calculating MHC Class II molecule binding score for each said sampled

segment by summing assigned values for each hydrophobic amino acid residue side chain present in said sampled amino acid residue segment; and (d) identifying at least one of said segments suitable for modification, based on the calculated MHC Class II molecule binding score for that segment, to change overall MHC Class II binding score for the peptide without substantially the reducing therapeutic utility of the peptide.

29. The method according to claim 28, wherein step (c) is carried out by using a Böhm scoring function modified to include 12-6 van der Waal's ligand-protein energy repulsive term and ligand conformational energy term by (1) providing a first data base of MHC Class II molecule models; (2) providing a second data base of allowed peptide backbones for said MHC Class II molecule models; (3) selecting a model from said first data base; (4) selecting an allowed peptide backbone from said second data base; (5) identifying amino acid residue side chains present in each sampled segment; (6) determining the binding affinity value for all side chains present in each sampled segment; and optionally (7) repeating steps (1) through (5) for each said model and each said backbone.

30. The method of claim 28 or 29, wherein the sampled amino acid residue segment is constituted by 13 amino acid residues.

31. The method of any of the claims 28 – 30, wherein consecutive sampled amino acid residue segments overlap by one to five amino acid residues.

32. The method of any of the claims 28 – 31, wherein 1 – 9 amino acid residues in any of the originally present T-cell epitope sequences are altered.

33. The method according to claim 32, wherein one amino acid residue in any of the originally present T-cell epitope sequences is altered.

34. The method of claim 32 or 33, wherein the alteration of the amino acid residues is substitution, deletion or addition of originally present amino acid(s) residue(s) by other amino acid residue(s) at specific position(s).

35. The method of claims 34, wherein additionally further alteration by substitution, deletion or addition is conducted to restore biological activity of said biological molecule.

36. An immunogenicly modified artificial protein selected from the group:

(i) Y – (L) – X, wherein Y is a cytokine and X, (L) is a molecule as defined above,

(ii) P – (L) – X, wherein P is a protein with unusual glycosylation moieties and X, (L) is a molecule as defined above,

(iii) A – (L) – X, wherein A is an immunoglobulin or a fragment thereof and X (L) is a molecule as defined above,

derived from a parent artificial protein having an amino acid sequence which is different from that of said parent artificial protein and exhibits reduced immunogenicity by a reduced number of T-cell epitopes relative to the parent fusion protein when exposed to the immune system of a given species, wherein said T-cell epitopes are peptide sequences able to bind to MCH class II molecule binding groups obtainable by a method as specified in any of the claims 28 – 35.

37. An artificial protein of claim 36, wherein at least A or X or Y or P is immunogenicly modified.

Fig.1

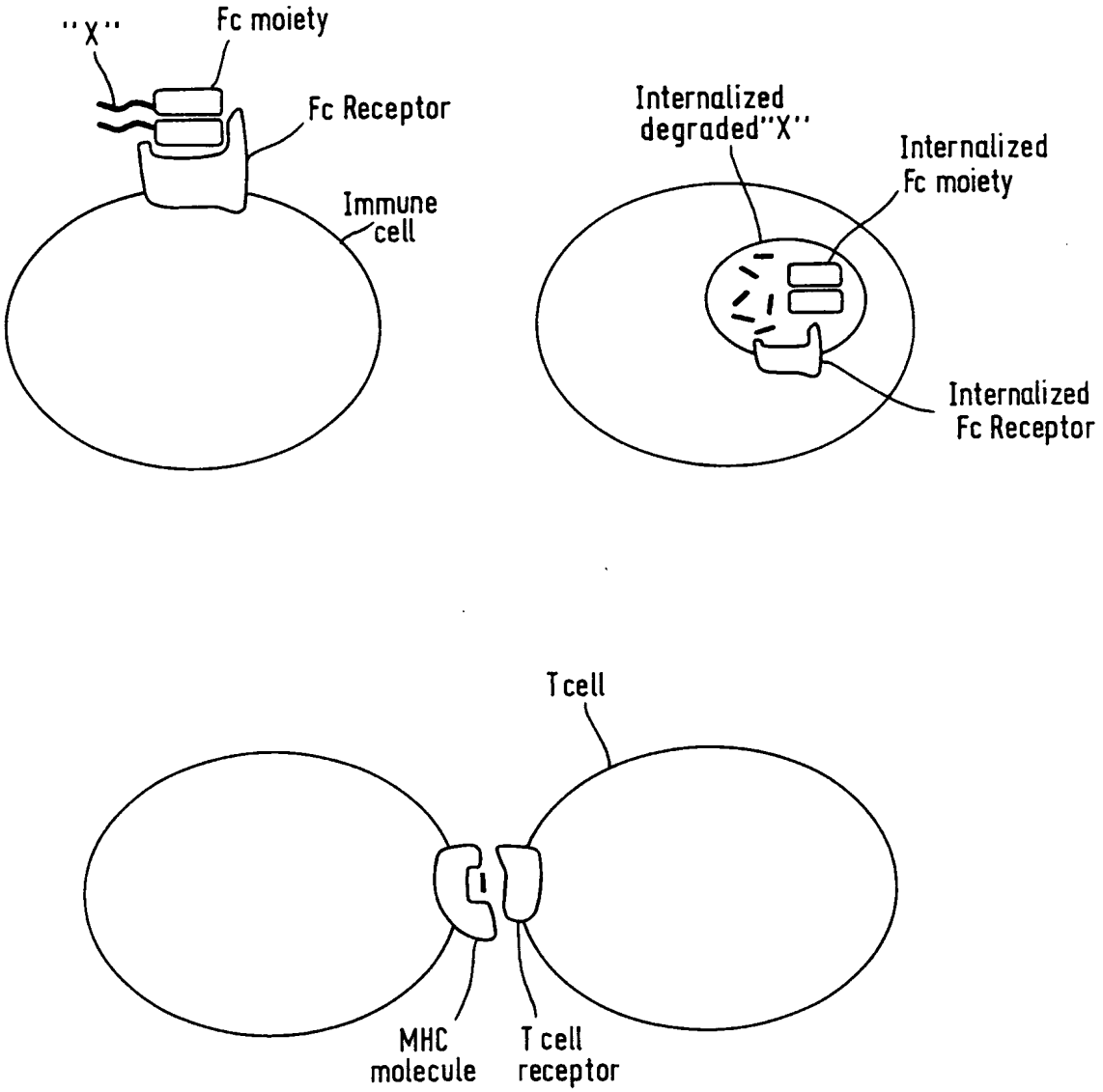

"X"
Fc moiety
Fc Receptor
Immune cell
Internalized degraded "X"
Internalized Fc moiety
Internalized Fc Receptor
Tcell
MHC molecule
T cell receptor

Fig.2

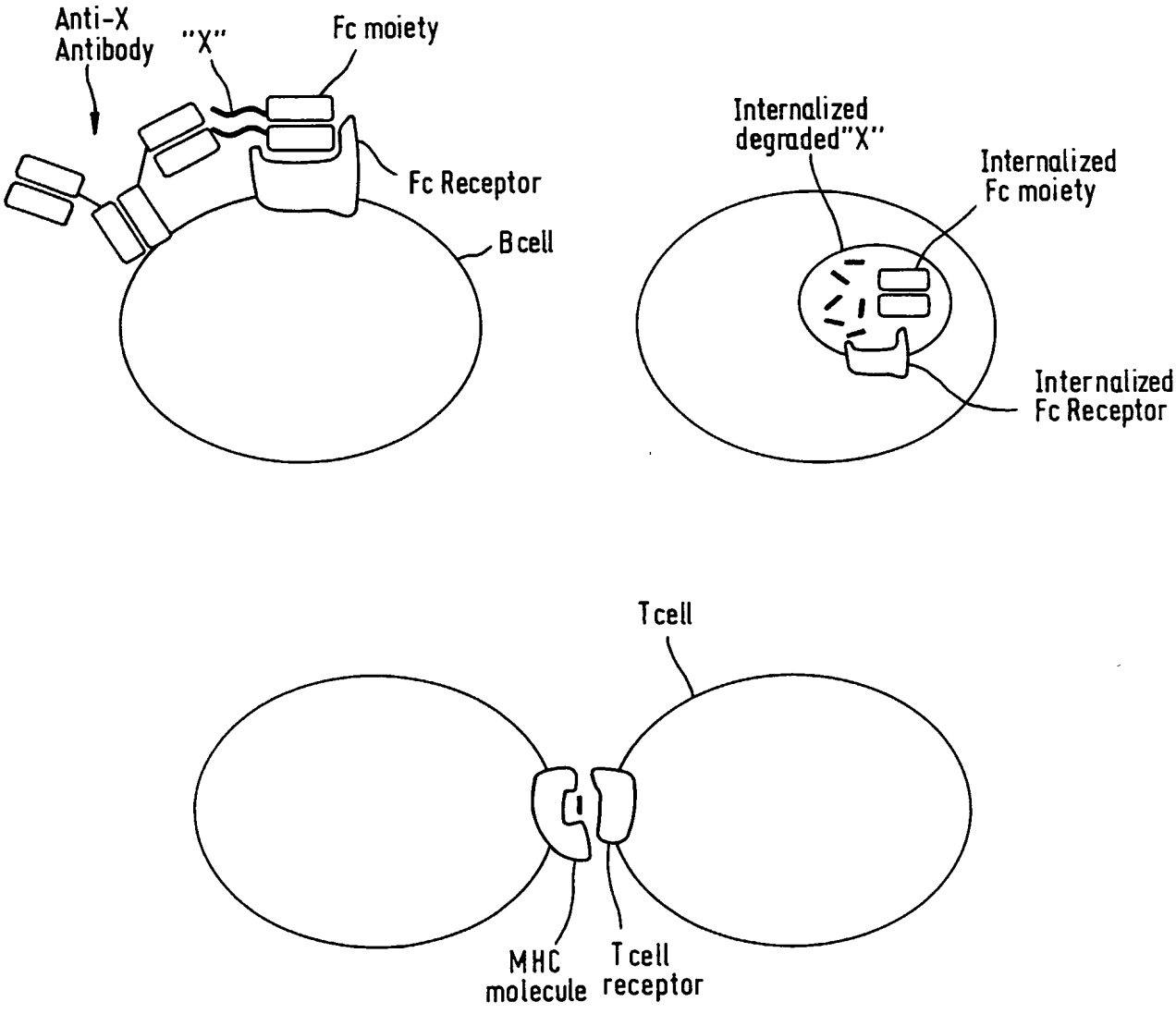

Anti-X
Antibody
"X"
Fc moiety
Fc Receptor
Bcell
Internalized
degraded"X"
Internalized
Fc moiety
Internalized
Fc Receptor
Tcell
MHC
molecule
Tcell
receptor

Fig.3

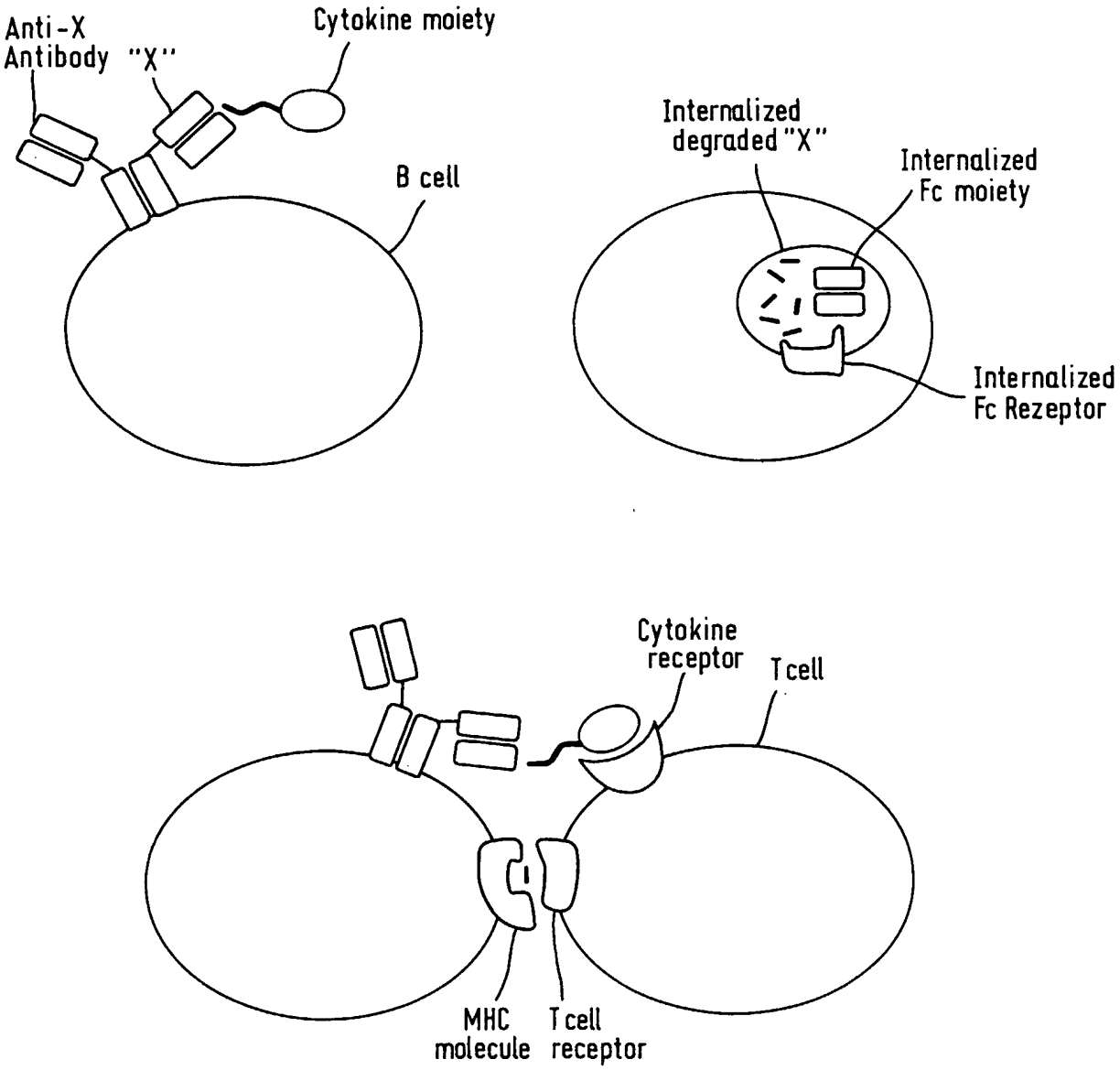
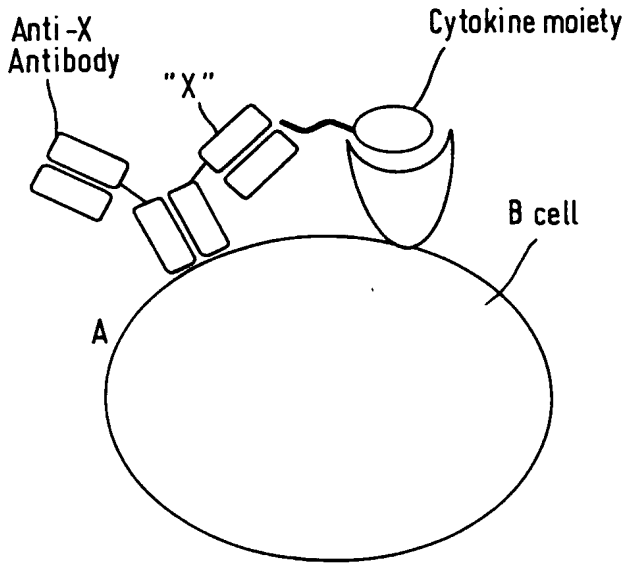

Anti-X
Antibody "X"
Cytokine moiety
B cell
Internalized
degraded "X"
Internalized
Fc moiety
Internalized
Fc Rezeptor
Cytokine
receptor
Tcell
MHC
molecule
Tcell
receptor

Fig.4



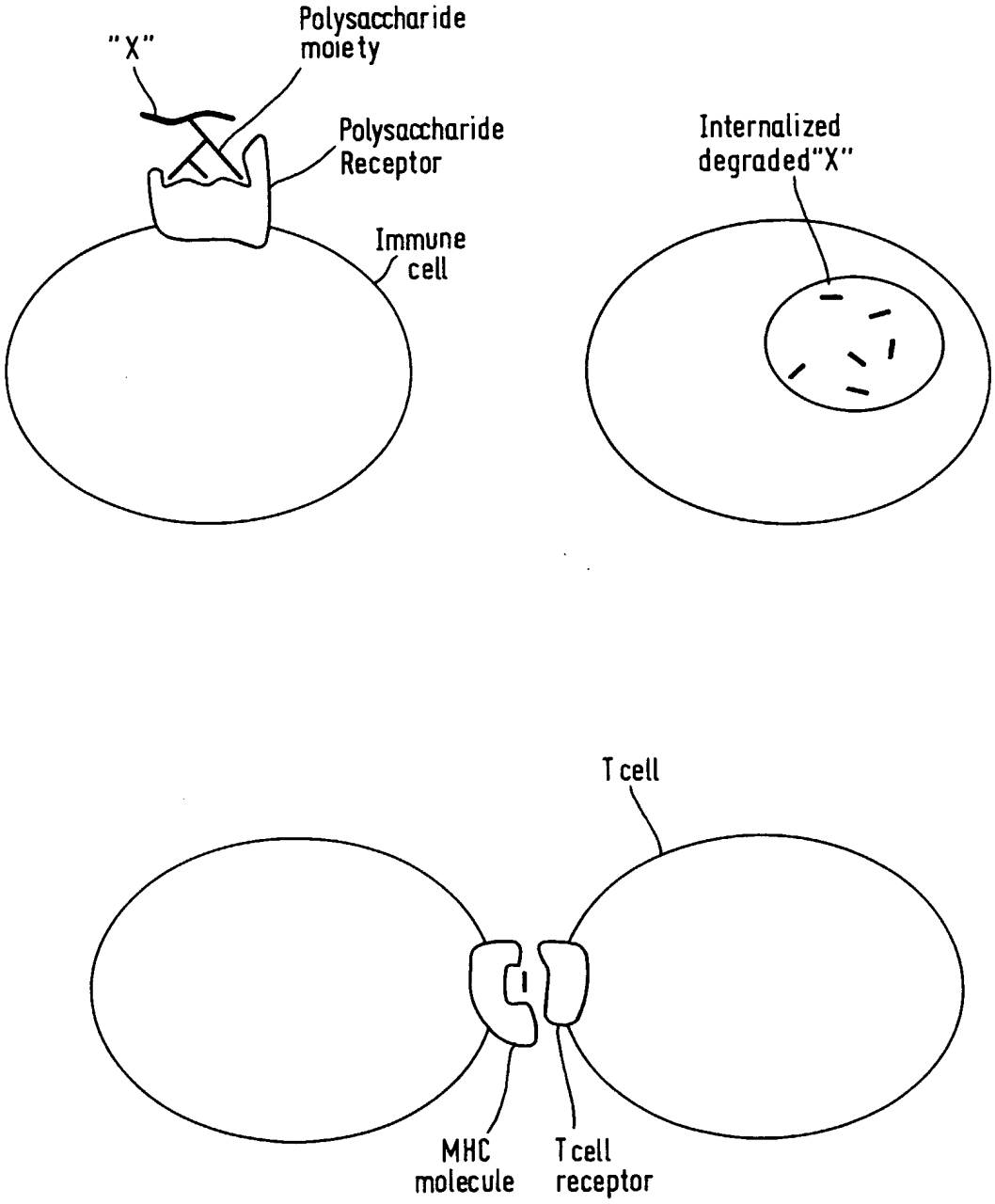
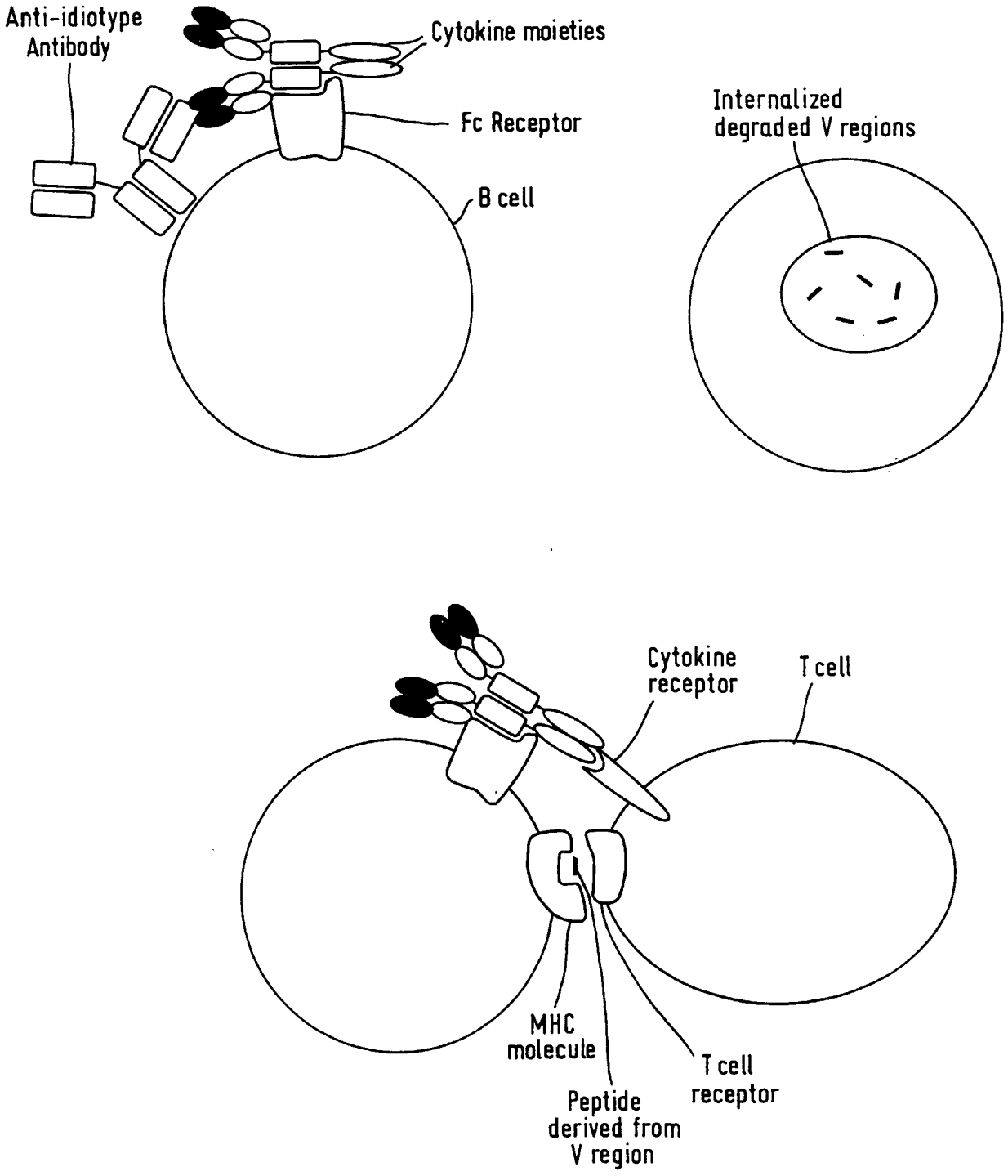

"X"
Polysaccharide moiety
Polysaccharide Receptor
Immune cell
Internalized degraded"X"
T cell
MHC molecule
T cell receptor

Fig.6


Anti-idiotype Antibody
Cytokine moieties
Fc Receptor
B cell
Internalized degraded V regions
Cytokine receptor
T cell
MHC molecule
Peptide derived from V region
T cell receptor